> *[F]rom a legal point of view there is nothing inherently unattainable about a prediction of future criminal conduct.*[1]

> *Electronic databases form the nervous system of contemporary criminal justice operations. In recent years, their breadth and influence have dramatically expanded. . . . The risk of error stemming from these databases is not slim. . . . Inaccuracies in expansive, interconnected collections of electronic information raise grave concerns for individual liberty.*[2]

INTRODUCTION

The Fourth Amendment requires "reasonable suspicion" to stop a suspect.[3] As a general matter, police officers develop this suspicion based on information they know or activities they observe. Suspicion is individualized to a particular person at a particular place.[4] Most reasonable suspicion cases involve police confronting unknown suspects engaged in observable suspicious activities.[5] Essentially, the reasonable suspicion doctrine is based on "small data"—discrete facts, limited information, and little knowledge about the suspect.[6]

But what happens if this small data suspicion is replaced by "big data" suspicion?[7] What if police can "know" personal information about the suspect by searching vast networked information sources? The rise of big data technologies offers a challenge to the traditional paradigm of Fourth Amendment law. With little effort, officers can now identify most unknown

[1] Schall v. Martin, 467 U.S. 253, 278 (1984).

[2] Herring v. United States, 555 U.S. 135, 155 (2009) (Ginsburg, J., dissenting).

[3] *See* Terry v. Ohio, 392 U.S. 1, 27 (1968).

[4] *See, e.g.*, United States v. Arvizu, 534 U.S. 266, 273 (2002) ("[W]e have said repeatedly that [courts] must look at the 'totality of the circumstances' of each case to see whether the detaining officer has a 'particularized and objective basis' for suspecting legal wrongdoing.").

[5] *See infra* Part I.

[6] "Small data," like "big data," has no set definition. Generally, small data is thought of as solving discrete questions with limited and structured data, and the data are generally controlled by one institution. *See generally* JULES J. BERMAN, PRINCIPLES OF BIG DATA: PREPARING, SHARING, AND ANALYZING COMPLEX INFORMATION 1-2 (2013).

[7] *See generally* Julie E. Cohen, *What Privacy Is For*, 126 HARV. L. REV. 1904, 1920-21 (2013) ("'Big Data' is shorthand for the combination of a technology and a process. The technology is a configuration of information-processing hardware capable of sifting, sorting, and interrogating vast quantities of data in very short times. The process involves mining the data for patterns, distilling the patterns into predictive analytics, and applying the analytics to new data. Together, the technology and the process comprise a technique for converting data flows into a particular, highly data-intensive type of knowledge."); Steve Lohr, *Amid the Flood, A Catchphrase Is Born*, N.Y. TIMES, Aug. 12, 2012, at BU3 [hereinafter Lohr, *Amid the Flood*] ("Big Data is a shorthand label that typically means applying the tools of artificial intelligence, like machine learning, to vast new troves of data beyond that captured in standard databases.").

suspects, not through their observations, but by accessing a web of information containing extensive personal data about suspects.[8] New data sources, including law enforcement databases, third-party records, and predictive analytics, combined with biometric or facial recognition software, allow officers access to information with just a few search queries.[9] At some point, inferences from this personal data (independent of the observation) may become sufficiently individualized and predictive to justify the seizure of a suspect. The question this Article poses is whether a Fourth Amendment stop can be predicated on the aggregation of specific and individualized, but otherwise noncriminal, factors.

For example, suppose police are investigating a series of robberies in a particular neighborhood. Arrest photos from a computerized database are uploaded in patrol cars. Facial recognition software scans people on the street.[10] Suddenly there is a match—police recognize a known robber in the targeted neighborhood. The suspect's personal information scrolls across the patrol car's computer screen—prior robbery arrests, prior robbery convictions, and a list of criminal associates also involved in robberies.[11] The officer then searches additional sources of third-party data, including the suspect's GPS location information for the last six hours or license plate records which tie the suspect to pawn shop trades close in time to prior robberies.[12] The police now have particularized, individualized suspicion about a man who is not doing anything overtly criminal. Or perhaps predictive software has already identified the man as a potential reoffender for this particular type of crime.[13] Or perhaps software has flagged the suspect's social media comments or other Internet postings that suggest planned criminal or gang

[8] *See infra* Part II.

[9] *See infra* Part II.

[10] *See infra* Part II; *see also Cop Car with Built-In Face Recognition and Predictive Policing Wins UK Award*, PRIVACYSOS.ORG (Apr. 4, 2013, 4:10 PM), http://privacysos.org/node/1016, *archived at* http://perma.cc/Y7BA-NTV2 (highlighting an example of technological advances in policing).

[11] All of this information is available through a National Crime Information Center (NCIC) search. *See National Crime Information Center*, FBI, http://www.fbi.gov/about-us/cjis/ncic (last visited Nov. 7, 2014), *archived at* http://perma.cc/P3CG-M5HF (explaining resources for finding information about criminals). This information is further available through police computers accessible in police cars and in police stations. *See National Crime Information Center Celebrates 40th Birthday*, GOV'T TECH. (Jan. 22, 2007), http://www.govtech.com/gt/articles/103437, *archived at* http://perma.cc/PDL7-JKS6 (discussing how NCIC records have helped law enforcement).

[12] *See infra* Part II.

[13] Local jurisdictions sometimes create their own "most wanted" lists of locally identified criminals. *See, e.g.*, *Most Wanted*, L.A. POLICE DEP'T, http://www.lapdonline.org/most_wanted (last visited Nov. 7, 2014), *archived at* http://perma.cc/9P5R-KZRG (showing a local jurisdiction's most wanted list).

activity.[14] Can this aggregation of individualized information be sufficient to justify interfering with a person's constitutional liberty?

This Article traces the consequences of a shift from "small data" reasonable suspicion, focused on specific, observable actions of unknown suspects, to a "big data" reality of an interconnected, information rich world of known suspects. With more specific information, police officers on the streets may have a stronger predictive sense about the likelihood that they are observing criminal activity.[15] This evolution, however, only hints at the promise of big data policing. The next phase will use existing predictive analytics to target suspects without any firsthand observation of criminal activity, relying instead on the accumulation of various data points.[16] Unknown suspects will become known to police because of the data left behind.[17] Software will use pattern-matching techniques[18] to identify individuals by sorting through information about millions of people contained in networked databases. This new reality simultaneously undermines the protection that reasonable suspicion provides against police stops and potentially transforms reasonable suspicion into a means of justifying those same stops.

This Article seeks to offer three contributions to the development of Fourth Amendment theory. First, it demonstrates that reasonable suspicion—as a small data doctrine—may become practically irrelevant in an era of big

[14] *See, e.g.*, Heather Kelly, *Police Embrace Social Media as Crime-Fighting Tool*, CNN (Aug. 30, 2012, 5:23 PM), http://www.cnn.com/2012/08/30/tech/social-media/fighting-crime-social-media, *archived at* http://perma.cc/D2LC-DEH8 (detailing ways police officers use social media to catch or thwart criminals).

[15] While this may protect some individuals who are not likely to be involved in criminal activity, it may also create additional burdens on those who are predicted to be involved in criminal activity. *See infra* Part IV.

[16] *See infra* Part II.

[17] *See* Paul Ohm, *Broken Promises of Privacy: Responding to the Surprising Failure of Anonymization*, 57 UCLA L. REV. 1701, 1716 (2010) (highlighting the difficulty of protecting the privacy of data subjects by anonymizing data); Paul M. Schwartz & Daniel J. Solove, *The PII Problem: Privacy and a New Concept of Personally Identifiable Information*, 86 N.Y.U. L. REV. 1814, 1877-78 (2011) (discussing identification of individuals from personally identifiable information found from data sources); Rebecca J. Rosen, *Stanford Researchers: It Is Trivially Easy to Match Metadata to Real People*, ATLANTIC (Dec. 24, 2013, 1:50 PM), http://www.theatlantic.com/technology/archive/2013/12/stanford-researchers-it-is-trivially-easy-to-match-metadata-to-real-people/282642/, *archived at* http://perma.cc/QFK5-6JUC (explaining the ease with which metadata can be matched with specific individuals).

[18] *See* Daniel J. Steinbock, *Data Matching, Data Mining, and Due Process*, 40 GA. L. REV. 1, 4 (2005) ("Data mining's computerized sifting of personal characteristics and behaviors (sometimes called 'pattern matching') is a more thorough, regular, and extensive version of criminal profiling, which has become both more widespread and more controversial in recent years."); Gareth Cook, *Software Helps Police Draw Crime Links*, BOS. GLOBE, July 17, 2003, at A1 (discussing how law enforcement officers are using databases as research tools).

data policing. Second, it examines the distortions of big data on police observation, investigation, and prediction, concluding that big data information will impact all major aspects of traditional policing. Third, it seeks to offer a solution to potential problems using the insights and value of big data itself to strengthen the existing reasonable suspicion standard.

Part I of this Article examines the development of Fourth Amendment law on reasonable suspicion. Much of this case law involves "unknown" suspects, such as when a police officer sees an individual on the street but does not know his or her identity. In these cases, reasonable suspicion necessarily derives from the suspect's observable actions. Most Fourth Amendment cases involving police–citizen encounters are of this "stranger" variety.[19] Thus, the reasonable suspicion test, as it evolved, required the police officer to articulate individualized, particularized suspicion to distinguish a stranger's suspicious actions from non-suspicious actions.[20] The resulting doctrine, created around actions, not individuals, makes sense within the context it arose (as presumably most officers would not know all of the potential criminals in their patrol areas).[21] The resulting reasonable suspicion test, however, becomes significantly distorted when officers have access to more individualized or predictive information about a suspect.

Part II of this Article addresses the rise of "big data" in criminal law enforcement. Law enforcement organizations are working to grow the scope, sophistication, and detail of their databases.[22] Agencies and their officers may now search national databases and gain instant access to the information.[23] Indeed, "data" is the new watchword in many smart-policing districts.[24]

[19] *See infra* Part I.

[20] *See* William J. Mertens, *The Fourth Amendment and the Control of Police Discretion*, 17 U. MICH. J.L. REFORM 551, 594-95 (1984) ("[T]he police must be able to justify singling out from the rest of humanity (or at least from the rest of the people in the general area) the particular individual whom they have stopped as somehow meriting this special attention.").

[21] This assumption is certainly true in large urban police districts, although it may hold less true for small towns or rural areas. As will be discussed later, "big data" in some ways turns big city policing into old-fashioned, small-town policing, with the benefits and drawbacks that come from that scale of police surveillance.

[22] *See infra* Part II.

[23] *See infra* Part II.

[24] *See* Nina Cope, *Intelligence Led Policing or Policing Led Intelligence?*, 44 BRIT. J. CRIMINOLOGY 188, 191 (2004) (discussing an operational structure for the organization of intelligence processes in police forces); Stephen Baxter, *Modest Gains in First Six Months of Santa Cruz's Predictive Police Program*, SANTA CRUZ SENTINEL (Feb. 26, 2012, 4:59 PM), http://www.santacruzsentinel.com/rss/ci_20050377, *archived at* http://perma.cc/KPM5-K634 (reporting on the success of a data algorithm used by the Santa Cruz Police Department); Carrie Kahn, *At LAPD, Predicting Crimes Before They Happen*, NPR (Nov. 26, 2011, 6:00 AM), http://www.npr.org/2011/11/26/142758000/at-lapd-predicting-crimes-before-they-happen, *archived at* http://perma.cc/P5JL-ZVWV (discussing how police use data to predict future crimes); Joel Rubin, *Stopping Crime Before It Starts*, L.A. TIMES (Aug. 21, 2010),

Crimes are recorded.[25] Criminals are cataloged.[26] Some jurisdictions record data about every police–citizen encounter, making both the person and justification for the stop (not necessarily even an arrest) instantly available to any officer.[27] Some jurisdictions have compiled "bad guy lists" identifying suspects in a neighborhood based on computer analysis of past actions and arrests.[28] In addition, law enforcement agencies increasingly rely on predictive algorithms to forecast individual recidivism and areas of likely criminal activity.[29]

Just as law enforcement agencies now collect and electronically analyze more personal data, so do private, third-party organizations.[30] These third-party entities are a familiar part of our daily lives. "Smartphones" record

http://articles.latimes.com/2010/aug/21/local/la-me-predictcrime-20100427-1, *archived at* http://perma.cc/N223-8J9K (suggesting that predictive policing that uses sophisticated data systems is the future of law enforcement).

[25] *Cf.* Andrew Guthrie Ferguson, *Crime Mapping and the Fourth Amendment: Redrawing 'High Crime Areas,'* 63 HASTINGS L.J. 179, 225-27 (2011) [hereinafter Ferguson, *Crime Mapping*] (discussing issues with recorded crime data).

[26] *Cf. id.* at 182 n.11.

[27] For example, in New York City, every stop-and-frisk is supposed to be recorded in an official UF-250 police report. *See* Bernard E. Harcourt & Tracey L. Meares, *Randomization and the Fourth Amendment*, 78 U. CHI. L. REV. 809, 862 n.210 (2011) ("According to the NYPD's Patrol Guide, a police officer who stops and frisks an individual must complete a UF-250 if a person is (1) stopped by force; (2) stopped and frisked or searched; (3) arrested; or (4) stopped and refuses to identify oneself. . . . In situations that fall outside these four contexts, a police officer may fill out a form if he or she desires to do so." (citation omitted)).

[28] *See* Stephen D. Mastrofski & James J. Willis, *Police Organization Continuity and Change: Into the Twenty-First Century*, *in* 39 CRIME & JUSTICE 55, 88 (Michael Tonry ed., 2010) ("Police now appear to rely more heavily on certain IT-based forms of surveillance—'database policing'—where officers use computers to 'patrol' massive data files (e.g., wanted lists) looking for 'hits' on information they possess on suspects."); Bryan Llenas, *Brave New World of "Predictive Policing" Raises Specter of High-Tech Racial Profiling*, FOX NEWS LATINO (Feb. 25, 2014), http://latino.foxnews.com/latino/news/2014/02/24/brave-new-world-predictive-policing-raises-specter-high-tech-racial-profiling/, *archived at* http://perma.cc/VG5W-WV93 ("[T]he Chicago Police Department, thanks to federal funding, is now helping to drive policing into territory previously only dreamed of in science fiction: The ability to essentially predict who will be the next perpetrator or the next victim of a crime."); Robert L. Mitchell, *Predictive Policing Gets Personal*, COMPUTERWORLD, (Oct. 24, 2013, 3:50 PM), http://www.computerworld.com/article/2486424/government-it/predictive-policing-gets-personal.html, *archived at* http://perma.cc/GDW5-B8JD ("Predictive policing is at the top of a lot of people's lists.").

[29] *See* Shima Baradaran & Frank L. McIntyre, *Predicting Violence*, 90 TEX. L. REV. 497, 507 (2012) (discussing how the majority of states detain or only conditionally release defendants determined to be dangerous); Andrew Guthrie Ferguson, *Predictive Policing and Reasonable Suspicion*, 62 EMORY L.J. 259, 265-69 (2012) [hereinafter Ferguson, *Predictive Policing*] (providing an overview of predictive policing).

[30] *See infra* Part II.

where we go.[31] Credit card companies record what we buy, and banks chronicle what we spend.[32] "OnStar" systems in cars catalog where and how fast we drive.[33] Phone records reflect our contacts and communications.[34] Internet searches reveal what we read and expose our interests.[35] Social media sites, such as Twitter and Facebook, even disclose what we think.[36] Currently, law enforcement officers may access many of these records without violating the Fourth Amendment, under the theory that there is no reasonable expectation of privacy in information knowingly revealed to third parties.[37] While certain statutory protections exist, most statutes include law enforcement exceptions,[38] and in any case, these private, commercial data aggregators have turned personal data into a commodity, available for purchase and analysis to anyone willing to pay.[39]

[31] *See, e.g.*, Eric Lichtblau, *Police Are Using Phone Tracking as Routine Tool*, N.Y. TIMES, Apr. 1, 2012, at 1.

[32] *Cf., e.g.*, Charles Duhigg, *Psst, You in Aisle 5*, N.Y. TIMES, Feb. 16, 2012 (Magazine), at 30 (discussing retail analytics).

[33] *See, e.g.*, Ned Potter, *Privacy Battles: OnStar Says GM May Record Car's Use, Even If You Cancel Service*, ABC NEWS (Sept. 26, 2011), http://abcnews.go.com/Technology/onstar-gm-privacy-terms-company-record-car-information/story?id=14581571, *archived at* http://perma.cc/VGN2-MJMZ.

[34] Phone companies record whom we call and even where we are located when we make those calls. *See, e.g.*, Noam Cohen, *It's Tracking Your Every Move, and You May Not Even Know*, N.Y. TIMES, Mar. 26, 2011, at A1.

[35] *See, e.g.*, Chloe Albanesius, *Facebook: Tracking Your Web Activity Even After You Log Out?*, PC MAG. (Sept. 26, 2011, 11:59 AM), http://www.pcmag.com/article2/0,2819,2393564,00.asp, *archived at* http://perma.cc/NWP2-DRXN; Robert Epstein, *Google's Gotcha*, U.S. NEWS & WORLD REP. (May 10, 2013, 12:15 PM), http://www.usnews.com/opinion/articles/2013/05/10/15-ways-google-monitors-you, *archived at* http://perma.cc/94V8-AUSX.

[36] *See generally* Noah Shachtman, *Exclusive: U.S. Spies Buy Stake in Firm That Monitors Blogs, Tweets*, WIRED (Oct. 19, 2009, 12:03 PM), http://www.wired.com/2009/10/exclusive-us-spies-buy-stake-in-twitter-blog-monitoring-firm/, *archived at* http://perma.cc/BZC6-SWAS (highlighting the intelligence value of social media posts).

[37] *See, e.g.*, Stephen E. Henderson, *Beyond the (Current) Fourth Amendment: Protecting Third-Party Information, Third Parties, and the Rest of Us Too*, 34 PEPP. L. REV. 975, 982-83 (2007) [hereinafter Henderson, *Beyond the (Current) Fourth Amendment*] (exploring the ways in which the third-party doctrine shortchanges privacy interests); Stephen E. Henderson, *Learning from All Fifty States: How to Apply the Fourth Amendment and Its State Analogs to Protect Third Party Information from Unreasonable Search*, 55 CATH. U. L. REV. 373, 376-79 (2006) [hereinafter Henderson, *Fifty States*] (describing the third-party doctrine); Orin S. Kerr, *The Case for the Third-Party Doctrine*, 107 MICH. L. REV. 561, 563 (2009) (same).

[38] Erin Murphy, *The Politics of Privacy in the Criminal Justice System: Information Disclosure, the Fourth Amendment, and Statutory Law Enforcement Exemptions*, 111 MICH. L. REV. 485, 487 & n.2 (2013) ("The United States Code currently contains over twenty separate statutes that restrict both the acquisition and release of covered information. . . . Yet across this remarkable diversity, there is one feature that all these statutes share in common: each contains a provision exempting law enforcement from its general terms.").

[39] *See infra* Part II.

The rise of "big data" means that this information is potentially available for use by law enforcement. In the same way that a drug store can predict that you will need a coupon this month because you bought a similar product last month,[40] the police will be able to anticipate that you will be selling drugs this week because you purchased an unusual number of mini–plastic bags last week.[41] Neither prediction is necessarily accurate, but both are based on individualized and particularized data that makes the prediction more likely.

Part III analyzes the intersection of big data and the current Fourth Amendment framework. The wrinkle of big data is that now officers are no longer dealing with "strangers." Even people unknown to officers can be identified and, with a few quick searches, revealed as a person with recognizable characteristics or about whom certain predictions can be made.[42] If officers view those individualized and particularized identifying characteristics—such as prior convictions, gang associations, and GPS coordinates near the scene of the crime—as suspicious, then otherwise innocent actions might create a predictive composite that satisfies the reasonable suspicion standard. In essence, reasonable suspicion will focus more on an individual's predictive likelihood of involvement in criminal activity than on an individual's actions.

Part III then looks at Fourth Amendment reasonable suspicion through three different lenses: (1) situations involving officers observing an ongoing crime, (2) situations involving officers investigating a past crime, and (3) situations involving officers predicting a future crime. Big data affects the analysis in each application, distorting the reasonable suspicion standard. Knowing who the suspect is and having more information (even innocent information) will allow officers to meet the reasonable suspicion threshold more easily because the information will be sufficiently individualized and particularized.

[40] *See* Duhigg, *supra* note 32, at 30; Rebecca Greenfield, *Facebook Now Knows What You're Buying at Drug Stores*, WIRE (Sept. 24, 2012, 11:49 AM), http://www.thewire.com/technology/2012/09/facebook-tracking-you-drug-store-now-too/57183/, *archived at* http://perma.cc/N5XH-QBA4; William F. Pewen, *Protecting Our Civil Rights in the Era of Digital Health*, ATLANTIC (Aug. 2, 2012, 11:09 AM), http://www.theatlantic.com/health/archive/2012/08/protecting-our-civil-rights-in-the-era-of-digital-health/260343/?single_page=true/, *archived at* http://perma.cc/8FB6-VERF.

[41] Mini–plastic bags (e.g., Ziploc bags) are used to package drugs sold on the street including cocaine, heroin, and marijuana. *See* United States v. Dingle, 114 F.3d 307, 309 (D.C. Cir. 1997) ("The government's narcotics expert testified that crack cocaine is typically packaged in small ziplock bags for street-level distribution."); United States v. Betts, 16 F.3d 748, 757 (7th Cir. 1994) (noting that pagers and Ziploc baggies are "hallmark paraphernalia" of drug distribution).

[42] *See infra* Part III.

Part IV assesses this new technological reality. Can the current reasonable suspicion doctrine adapt? Should it? What are the possible benefits or dangers of big data reasonable suspicion? Using big data may help reduce the negative consequences of traditional policing techniques, but at the same time may create a whole new set of concerns. This section evaluates the tradeoffs of big data as applied to the Fourth Amendment.

Part V offers a few solutions to the problem presented by the big data distortions of Fourth Amendment doctrine. This Article suggests that the nature of big data itself might provide a means of strengthening the reasonable suspicion standard. If big data resources are used to tip the scales of reasonable suspicion in favor of law enforcement, then courts should require a higher level of detail and correlation using the insights and capabilities of big data. This requirement would involve precise statistical analysis, geospatial analysis, temporal analysis, and link analysis of the data. Big data can provide information about a person on a generalized or granular scale, and the latter should be required. The power of big data allows investigators to go deep into the data and make sure that the information is as tightly correlated as possible. In this way, a big data–infused reasonable suspicion standard will do what the reasonable suspicion requirement was always supposed to do—distinguish the criminal from the noncriminal in a manner that balances the need for effective law enforcement with a measure of personal liberty.

I. REASONABLE SUSPICION: A SMALL DATA DOCTRINE

The Fourth Amendment serves as a constitutional barrier, protecting individuals from unreasonable police intrusion.[43] On the street, the police may not constitutionally stop, seize, or search individuals without the requisite legal justification.[44] To seize a person temporarily, a police officer

[43] U.S. CONST. amend. IV; Dunaway v. New York, 442 U.S. 200, 213 (1979) ("Hostility to seizures based on mere suspicion was a prime motivation for the adoption of the Fourth Amendment, and decisions immediately after its adoption affirmed that 'common rumor or report, suspicion, or even "strong reason to suspect" was not adequate to support a warrant for arrest.'" (citing Henry v. United States, 361 U.S. 98, 101 (1959))); Almeida-Sanchez v. United States, 413 U.S. 266, 273 (1973) ("The needs of law enforcement stand in constant tension with the Constitution's protections of the individual against certain exercises of official power. It is precisely the predictability of these pressures that counsels a resolute loyalty to constitutional safeguards.").

[44] *See* Brown v. Texas, 443 U.S. 47, 51 (1979) ("A central concern in balancing these competing considerations in a variety of settings has been to assure that an individual's reasonable expectation of privacy is not subject to arbitrary invasions solely at the unfettered discretion of officers in the field.").

and the Stored Communications Act,[256] but the protection lapses quickly.[257] Finally, telephone records are subject to protection through the Telephone Records and Privacy Protection Act of 2006,[258] but they too can be accessed by police if the evidence is relevant based on "specific and articulable" facts.[259]

In addition to constitutional and statutory protections, certain consumer guidelines established by companies promise to keep information private.[260] Yet most major commercial entities—including Internet search companies, online retailers, and social media platforms—collect data to monetize it.[261] In fact, many businesses, including big-name companies like Google, Microsoft, Yahoo!, and Facebook, are financially successful, in part, because of their ability to sell targeted advertising using user data.[262] These economic incentives, combined with a willingness to assist law enforcement as good corporate citizens, means that most third-party information is not well-protected from government access.

III. BIG DATA AND REASONABLE SUSPICION ON THE STREETS

What happens when a doctrine built on small data becomes overwhelmed by big data? What happens when previously unknown suspects can become known with a few quick search queries? Police and courts will soon confront this new reality as officers come to use existing facial recognition or biometric technology and networked databases to obtain individualized and particularized information about a suspect. Courts will confront additional questions as these technologies become more sophisticated, mobile, and reliant on predictive analytics.

[256] 18 U.S.C. §§ 2701–2712 (2012).

[257] *Compare* 18 U.S.C. §§ 2511, 2516, 2518 (2012) (describing the heightened requirements for obtaining real time communications), *with id.* § 2703(a) (setting out the lower standards for obtaining a court order for stored communications).

[258] 18 U.S.C. § 1039 (2012).

[259] *Id.* § 2703(c)–(d).

[260] *See, e.g., Microsoft.com Privacy Statement*, MICROSOFT, http://privacy.microsoft.com/en-us/default.mspx (last updated Aug. 2013), *archived at* http://perma.cc/F96M-8FUH; *Privacy Policy*, GOOGLE, http://www.google.com/privacy (last modified Mar. 31, 2014), *archived at* http://perma.cc/5FL4-NEHK;

[261] *See supra* Section II.B.2.

[262] *See, e.g.*, Rupert Neate & Rowena Mason, *Networking Site Cashes in on Friends*, TELEGRAPH (Jan. 31, 2009), http://www.telegraph.co.uk/finance/newsbysector/mediatechnologyandtelecoms/4413483/Networking-site-cashes-in-on-friends.html, *archived at* http://perma.cc/CBF6-R5N9 (reporting Facebook's move to monetize its collection of personal user information by allowing advertisers to target Facebook users selectively).

This Part studies this intersection of technology and doctrine through three different lenses—observation, investigation, and prediction—mirroring the most common types of police work. Police officers regularly observe ongoing criminal activity, investigate past criminal activity, and predict future criminal activity. The impact of "big data suspicion" will be different depending on the type of police activity at issue.

A. *Observation of Ongoing or Imminent Crimes*

Consider a modern day *Terry v. Ohio* situation. Detective McFadden is patrolling the street. He observes John Terry and, using facial recognition technology, identifies him and begins to investigate using big data. Detective McFadden learns through a database search that Terry has a prior criminal record, including a couple of convictions and a number of arrests.[263] McFadden learns, through pattern–matching links, that Terry is an associate (a "hanger on") of a notorious, violent local gangster—Billy Cox—who had been charged with several murders.[264] McFadden also learns that Terry has a substance abuse problem and is addicted to drugs.[265] These factors—all true, but unknown to the real Detective McFadden—are individualized and particularized to Terry. Alone, they may not constitute reasonable suspicion that Terry is committing or about to commit a particular crime. But in conjunction with Terry's observed actions of pacing outside a store with two associates, the information makes the reasonable suspicion finding easier and, likely, more reliable.

In observation cases, by using mobile facial recognition to identify the suspect, the officer now can turn any unknown suspect into a known suspect and can search for information that might justify reasonable suspicion. This change allows the officer to review traditional data sources known to law enforcement, including prior criminal history, arrests, addresses, gang associations, known associates, and even concealed weapons permits. Perhaps this individual is on a local "most wanted" list or a watch list as someone who has already been identified as being trouble in the neighborhood.[266] Perhaps

[263] Louis Stokes, *Representing John W. Terry*, 72 ST. JOHN'S L. REV. 727, 728-29 (1998) (discussing the facts of the *Terry* case); *see also* Terry v. Ohio *30 Years Later*, *supra* note 48, app.B at 1523 (reporting the sentencing judge as describing Terry as "a man who has from December 30, 1948, to the present time, be[en] consistently involved in difficulties with the law").

[264] Stokes, *supra* note 263, at 728-29.

[265] *Id.* at 727.

[266] *See* Slobogin, *Government Data Mining*, *supra* note 138, at 322 ("Match-driven data mining programs are designed to determine whether a particular individual has already been identified as a 'person of interest.' In other words, the goal here is not to find out more about a suspect, but rather to determine whether a particular person is a known suspect." (emphasis omitted)).

his height, weight, race, hairstyle, facial hair, or other distinguishing marks match a robbery suspect. This traditional law enforcement information might also now include data from automatic license plate readers, digitally archived surveillance video, and intelligence reports created and maintained by police. Even this limited information may be—as a constitutional matter—enough for an officer to stop the suspect. If, for example, the suspect had an extensive history of commercial robberies, or if license plate data connected him to prior robberies in the area, this information might well constitute reasonable suspicion that the suspect was going to commit a robbery.

Additional big data innovations may also assist the police. For example, the New York Police Department (NYPD) has unveiled the Domain Awareness System (DAS) in partnership with Microsoft.[267] This technology allows an officer to observe, through video surveillance or automated license plate readers, the location of a suspect prior to the initial observation:

> DAS is capable of rapidly blending and analyzing realtime data gathered from roughly 3,000 civic closed-circuit cameras, 911 call recordings, and license plate readers . . . as well as historical crime reports. Now the NYPD can do things like track a vehicle and instantly determine nearly everywhere it's been for the past few days or weeks; instantly access a suspect's arrest record, and all the 911 calls related to a particular crime; [and] map criminal history to geospatially and chronologically reveal crime patterns . . . .[268]

Thus, the officer could determine whether the suspect had just arrived with a getaway driver, had been casing the store, or had merely been doing non-criminal errands all morning.[269] These patterns may well affect whether an officer has reasonable suspicion that a suspect is about to commit a crime.

[267] *See* Press Release, Microsoft, New York City Police Department and Microsoft Partner to Bring Real-Time Crime Prevention and Counterterrorism Technology Solution to Global Law Enforcement Agencies (Aug. 8, 2012), *available at* http://news.microsoft.com/2012/08/08/new-york-city-police-department-and-microsoft-partner-to-bring-real-time-crime-prevention-and-counter terrorism-technology-solution-to-global-law-enforcement-agencies.

[268] Douglas Page, *Crime Fighting's Next Big Deal*, OFFICER.COM (Sept. 4, 2012), http://www.officer.com/article/10773317/crime-fightings-next-big-deal, *archived at* http://perma.cc/YTF5-A2UC; *see also* Michael Endler, *NYPD, Microsoft Push Big Data Policing Into Spotlight*, INFO. WK. (Aug. 20, 2012), http://www.informationweek.com/security/privacy/nypd-microsoft-push-big-data-policing-in/240005838, *archived at* http://perma.cc/DK97-7HMD (describing how DAS could lead to earlier apprehension of criminals).

[269] Somini Sengupta, *Privacy Fears as Surveillance Grows in Cities*, N.Y. TIMES, Oct. 13, 2013, at A1 (pointing out that big data-driven policing in Oakland, California, could help separate innocent actions from criminal activity).

For a second level of inquiry, imagine the police officer uses networked databases owned by third parties to discover personal information about a suspect. This data might include credit information, financial records, credit card activity, employment, past addresses and telephone numbers, names and addresses of family members, neighbors' addresses and telephone numbers, business associates, make, model, and color of registered vehicles, social security numbers, dates of birth, bankruptcies, liens and judgments, and GPS locational data. While access to some of these data would usually require particular legal authorization, law enforcement can circumvent statutes restricting direct access by instead using "fourth-party" commercial aggregators.[270] Such personalized information will allow an officer to develop a more individualized picture of a suspect. While generally unemployment, credit card debt, and bankruptcy are not indicia of criminal activity, when viewed in conjunction with suspicious action in front of an expensive jewelry store, however, a personal financial crisis might be relevant to the totality of circumstances. Further, accurate GPS data tying the suspect to a prior robbery or to a pawnshop might lead to reasonable suspicion. Even the otherwise innocent purchase of a wool cap or ski mask at Walmart might tip a seasonal purchase into reasonable suspicion.

Finally, imagine if law enforcement could access the suspect's social media data.[271] Search queries, Facebook and Twitter posts, YouTube videos, emails, texts, and similar communications are all available to third-party providers—if not publically available. While personal content is usually statutorily (or commercially) protected, it is generally not constitutionally protected. This mosaic of personal information might well provide individualized facts necessary to make the police officer's suspicion reasonable.[272] For example, a suspect's admission of financial difficulties or photograph displaying the fruits of the crime through social media could appropriately be added to the totality of circumstances.

With each level of search, officers can access additional individualized and particularized facts that, when viewed within the totality of circumstances, help justify the officer's stop of a suspect. The effect is that additional personalized information encourages a finding of reasonable suspicion. A

[270] *See* Simmons, *supra* note 173, at 990-92 (describing commercial data acquisitions by the government).

[271] *See* Joseph Goldstein & J. David Goodman, *Seeking Clues to Gangs and Crimes, Detectives Follow Internet Rap Videos*, N.Y. TIMES, Jan. 8, 2014, at A20 ("Directed by prosecutors to build evidence that individual shootings are part of larger criminal conspiracies, officers are listening to local rappers for a better sense of the hierarchy of the streets. 'You really have to listen to the songs because they're talking about ongoing violence.'").

[272] *Cf. id.* (highlighting police use of social media to gain insight into criminal conspiracies).

trip to a pawnshop could indicate a person is selling stolen goods—or is merely poor enough to have to sell belongings at a steep discount. A photograph of jewelry could be an admission of theft or could simply be a photograph of jewelry. Yet in a criminal investigation, the inferences of suspicion are easy to develop and, against a low legal threshold, easy to meet.

Of course, suspicious facts must be connected with a suspected crime. It would not be relevant if the searches revealed a pattern of domestic violence crimes, unrelated to robbery. It would also not be relevant if the information was not directly connected to the suspect. Being a friend of a friend of a known robber is a fact, but not one that should influence the constitutional calculus. But, as long as the data are connected to both the suspected criminal activity and the suspected criminal, it would likely be persuasive in evaluating reasonable suspicion in observation cases.

B. *Investigation of Completed Crimes*

Many crimes occur without direct police observation, and police must investigate the crime to identify the perpetrator. Reasonable suspicion is still relevant in investigating past crimes (assuming the information available does not rise to the higher level of probable cause).[273] In *Hensley*, the Supreme Court set out the standard for investigating past crimes based on reasonable suspicion:

> The precise limits on investigatory stops to investigate past criminal activity are more difficult to define. The proper way to identify the limits is to apply the same test already used to identify the proper bounds of intrusions that further investigations of imminent or ongoing crimes. That test, which is grounded in the standard of reasonableness embodied in the Fourth Amendment, balances the nature and quality of the intrusion on personal security against the importance of the governmental interests alleged to justify the intrusion. When this balancing test is applied to stops to investigate past crimes, we think that probable cause to arrest need not always be required.[274]

[273] Illinois v. Gates, 462 U.S. 213, 241 (1983) ("[P]robable cause deals with probabilities." (internal quotation marks omitted)). The impact of big data on probable cause is a separate subject beyond the scope of this Article.

[274] United States v. Hensley, 469 U.S. 221, 228 (1985) (citations omitted); *see also id.* at 227 ("This is the first case we have addressed in which police stopped a person because they suspected he was involved in a completed crime. In our previous decisions involving investigatory stops on less than probable cause, police stopped or seized a person because they suspected he was about to commit a crime, or was committing a crime at the moment of the stop." (citation omitted)); *id.* ("We do not agree . . . that our prior opinions contemplate an inflexible rule that precludes police

While acknowledging that courts might balance these interests differently when investigating a past, completed crime—as opposed to an ongoing crime[275]—the Supreme Court still held that "the ability to briefly stop that person, ask questions, or check identification in the absence of probable cause promotes the strong government interest in solving crimes and bringing offenders to justice."[276] By adopting a reasonable suspicion test for investigation of past crimes, the Court gave police the flexibility to stop suspects based on this lower threshold of suspicion.[277]

As in observation cases, the primary use of big data would be to identify unknown perpetrators for arrest and prosecution. As one security expert explained, "[i]magine the ability to instantly take a security camera photograph from a bank robbery and match it using a facial recognition algorithm to a photograph in an out-of-state motor vehicle database, and then to link that person's name to a mobile phone from a private-sector marketing database."[278] Already, police have relied on similar linkages of networked information in more run-of-the-mill cases.[279] With new search technology, disparate pieces of data are compiled to link, match, and identify a suspect through pattern matching techniques. This can be done not only with a name, address, or license plate, but also with a particular modus operandi.[280]

from stopping persons they suspect of past criminal activity unless they have probable cause for arrest. To the extent previous opinions have addressed the issue at all, they have suggested that some investigative stops based on a reasonable suspicion of past criminal activity could withstand Fourth Amendment scrutiny.").

[275] *See id.* at 228-29 ("The factors in the balance may be somewhat different when a stop to investigate past criminal activity is involved rather than a stop to investigate ongoing criminal conduct.").

[276] *Id.* at 229.

[277] United States v. Place, 462 U.S. 696, 702 (1983) (allowing stops "when the officer has reasonable, articulable suspicion that the person *has been*, is, or is about to be engaged in criminal activity." (emphasis added)); Florida v. Royer, 460 U.S. 491, 498 (1983) (allowing certain seizures "if there is articulable suspicion that a person *has committed* or is about to commit a crime" (emphasis added)); United States v. Cortez, 449 U.S. 411, 417 n.2 (1981) ("Of course, an officer may stop and question a person if there are reasonable grounds to believe that person is wanted for past criminal conduct.").

[278] Page, *supra* note 268.

[279] *See* Mark Ward, *Crime Fighting with Big Data Weapons*, BBC (Mar. 18, 2014, 2:35 AM), http://www.bbc.com/news/business-26520013, *archived at* http://perma.cc/4ETS-GKDF; *see also* Neal Ungerleider, *This Small City's Police Department Builds an App, Nabs Big Data to Find and Fight Bad Guys*, FAST COMPANY (Mar. 26, 2014, 9:00 AM), http://www.fastcompany.com/3027641/this-small-citys-police-department-builds-an-app-nabs-big-data-to-find-and-fight-bad-guys, *archived at* http://perma.cc/7Z7H-5TNP.

[280] Taipale, *supra* note 184, at 21 ("The popular view of investigation in law enforcement is that there must first be a specific crime and that law enforcement then follows particularized clues or suspicions after the fact. In reality, investigators often look for criminal patterns or hypothetical suspects in order to anticipate future crime. For example, investigators may use pattern recogni-

This information, specific to a person and particularized to a crime, meets both requirements needed to establish reasonable suspicion.

The value of big data to reasonable suspicion investigations is probably greater than its value to observation cases, because police have time to surmount the "legal process" requirements necessary to obtain third-party information.[281] With an official request, a court order, or a subpoena (let alone a warrant or grand jury subpoena), law enforcement officers can obtain most third-party data if doing so in furtherance of a criminal investigation.[282]

Software can isolate patterns and identify suspects through existing public and private data in novel ways. One fascinating example of big data sleuthing arose out of the investigation of a major Swedish armed robbery of millions of dollars.[283] Police assumed that, to disguise their plot, the thieves must have used prepaid disposable phones. Data analysts then searched through the list of all prepaid disposable phones in the area looking for "a set of phones that stayed within their own miniature network."[284] Police analysts found a single set of phones that only communicated with each other, did so only for a few weeks leading up to the heist, and then went silent after the robbery. Identifying this network allowed police to solve the case. Police traced the phones to specific cell tower locations corresponding with the robbers' locations before, during, and after the robbery.[285] In fact, once police knew the numbers, they could track location-by-location exactly where the robbers had been. When police identified one person who had purchased the phones, they were able to determine how the crime occurred and the location of the thieves at all times.[286]

Major police departments, as well as the FBI, have adopted this type of pattern matching investigation technique.[287] In child abduction cases,

tion strategies to develop modus operandi ('MO') or behavioral profiles, which in turn may lead either to specific suspects (profiling as identifying pattern) or to crime prevention strategies (profiling as predictor of future crime, resulting, for example, in stakeouts of particular places, likely victims, or potential perpetrators).").

[281] *See supra* Section II.D (discussing the statutory requirements of court orders for some private information).

[282] *See supra* Section II.D (noting the ease with which law enforcement may access records that are protected by statute).

[283] EVAN RATLIFF, LIFTED ch. 5–9 (Kindle Singles ed. 2011), *available at* https://www.atavist.com/stories/lifted/ (describing the investigation of the heist).

[284] *Id.* at ch. 12.

[285] *Id.*

[286] *Id.* at ch. 13.

[287] *See* Josh Richman & Angela Woodall, *Around the Bay Area, You're Being Watched*, CONTRA COSTA TIMES (June 30, 2013, 1:29 AM), http://www.contracostatimes.com/News/ci_23569173/Around-the-Bay-Area-youre-being, *archived at* http://perma.cc/V8UE-2TJR ("[I]t's not just the

"Amber alerts" have led to quick reviews of license plate reader databases. By searching the location of a car, police can determine the likely route of the suspect.[288] In gang cases, recordings of gunshots have helped map out areas of contested gang turf.[289]

Returning to the robbery example, imagine that a particular jewelry store was robbed by an unknown suspect. Police officers have a video still from the robbery that does not allow for a facial recognition match. The photo, however, clearly shows a neck tattoo, and officers obtain a partial description of the getaway car. Running a search for the tattoo against a database might narrow the list of suspects. Comparing the narrowed list with owners of a particular type of car might further limit the list of suspects. Looking at the remaining suspects' associates, movements, or even bank deposits, credit card expenditures, or social media comments might again tighten the search. The result is that big data can help identify the suspect with a few search queries. While these data might not be enough to get an arrest warrant, they would likely provide the reasonable suspicion needed to stop and investigate the suspect.[290]

C. *Predicting Crimes*

Unlike observation or investigation cases, reasonable suspicion based on prediction remains the stuff of science fiction. Police have begun to predict areas of heightened criminal activity,[291] and may predict likely troublemakers

National Security Agency secretly vacuuming up your personal data. Local police agencies are increasingly adopting Big Data technologies . . . ."); *cf.* Charles Piller & Eric Lichtblau, *FBI Plans to Fight Terror with High-Tech Arsenal*, L.A. TIMES, July 29, 2002, at A1 ("By Sept. 11, 2011, the FBI hopes to use artificial-intelligence software to predict acts of terrorism the way the telepathic 'precogs' in the movie 'Minority Report' foresee murders before they take place.").

[288] Lochner, *supra* note 199, at 225 ("[T]he Automated License Plate Recognition system, store[s] license plate numbers of the innocent and guilty so the database can be mined during Amber Alerts or for leads in cases.").

[289] *See* Christopher Benjamin, Note, *Shot Spotter and FaceIt: The Tools of Mass Monitoring*, UCLA J.L. & TECH., Spring 2002, art. 2, at 6 (describing a system by which automated phone calls help find the location of gunfire).

[290] *See* Cook, *supra* note 18 ("The Boston Police Department is rolling out a powerful new computer program built to find hidden connections among people and events almost instantly, allowing detectives to investigate murders, rapes, and other crimes far faster than they can today."); *see also* Yang Xiang et al., *Visualizing Criminal Relationships: Comparison of a Hyperbolic Tree and a Hierarchical List*, 41 DECISION SUPPORT SYS. 69, 75-77 (2005) (describing how a tool known as COPLINK Criminal Relationship Visualizer links co-occurring events and characteristics).

[291] Ferguson, *Predictive Policing*, *supra* note 29, at 312-13; Paul Bowers, *Predictive Policing Arrives in Charleston*, CHARLESTON CITY PAPER (June 27, 2012), http://www.charlestoncitypaper.com/charleston/predictive-policing-arrives-in-charleston/Content?oid=4101684, *archived at* http://perma.cc/JWL7-35TD (discussing the use of predictive analytics to reduce armed robberies in Charleston, South Carolina).

involved in criminal enterprise through an unofficial "bad guy list," but predictive analytics cannot yet tell police whom to stop for a crime not yet committed. To be clear, these are prediction-based stops where no crime has occurred and no crime is observed.

Yet big data invites provocative questions about whether such predictive tips should factor into the reasonable suspicion calculus. For example, if a drug distribution gang is run by a tight-knit family or neighborhood organization, such that the pattern for several years has been that when one family member is arrested, another cousin or brother takes their place, then why can we not predict who will be the next member of the gang?[292] If burglaries are contagious in part because the same gang of burglars commits similar crimes, and police identify one burglar, why should we not target a burglar's associates as likely suspects for future burglaries?[293] In these cases, police could show specific and articulable facts indicating that a particular person is likely to participate in ongoing criminal activity (e.g., drug dealing or burglaries).[294] Because the criminal enterprise is ongoing, the *Terry* standard might well apply, and police could try to stop and investigate would-be members of these criminal organizations if they were observed doing anything that might suggest drug dealing or burglary.

The questions get harder when no ongoing criminal enterprise exists, yet the same predictive logic holds. In Chicago, analysts have identified young people at greater risk of being involved in gun violence.[295] Researchers

[292] *See* STEVEN D. LEVITT & STEPHEN J. DUBNER, FREAKONOMICS: A ROGUE ECONOMIST EXPLORES THE HIDDEN SIDE OF EVERYTHING 89-114 (2005) (discussing the economics and social relationships of the drug trade in the famous chapter "Why Do Drug Dealers Still Live with Their Moms?").

[293] *See* Wim Bernasco, *Them Again?: Same-Offender Involvement in Repeat and Near Repeat Burglaries*, 5 EUR. J. CRIMINOLOGY 411, 423-25 (2008) ("[B]oth repeat burglaries and near repeat burglaries are much more likely to involve the same offender than are spatially or temporally unrelated burglaries."); Bowers & Johnson, *supra* note 229, at 13 (discussing how features of an offender's modus operandi, like spatial and temporal preferences, can be used to identify crimes carried out by a particular network of offenders).

[294] Domestic violence also presents a possible predictive environment for crime. *See* Joseph Goldstein, *Police Take on Family Violence to Avert Death*, N.Y. TIMES, July 25, 2013, at A1 ("[T]he officers assigned to the domestic violence unit make a total of 70,000 precautionary visits a year to the households with past episodes. Each precinct station house also maintains a 'high propensity' list of a dozen or so households that get special attention because they are believed to be most at risk of further violence.").

[295] *See* Andrew Papachristos, Tracey L. Meares & Jeffrey Fagan, *Attention Felons: Evaluating Project Safe Neighborhoods in Chicago* 4 J. EMPIRICAL LEGAL STUD. 223, 229-33 (describing Chicago's program to identify and address likely perpetrators and victims of gun violence); *see also* TRACEY MEARES, ANDREW V. PAPACHRISTOS & JEFFREY FAGAN, HOMICIDE AND GUN VIOLENCE IN CHICAGO: EVALUATION AND SUMMARY OF THE PROJECT SAFE NEIGHBORHOODS PROGRAM 1 (2009), *available at* http://www.psnchicago.org/PDFs/2009-PSN-Research-Brief_v2.pdf ("Data analysis

can predict their likelihood of being a victim or perpetrator of gun violence using big data metrics, including place of residence, social associations (e.g., past experience with victims of gun violence and gang connections), and age.[296] Assuming the accuracy of these data, could police target these individuals as part of a predictive stop strategy?[297] In fact, the Chicago Police Department appears to have adopted this predictive logic in its intervention program. As described by the *New York Times*,

> [i]n recent months, as many as 400 officers a day, working overtime, have been dispatched to just 20 small zones deemed the city's most dangerous. The police say they are tamping down retaliatory shootings between gang factions by using a comprehensive analysis of the city's tens of thousands of suspected gang members, the turf they claim and their rivalries. The police are also focusing on more than 400 people they have identified as having associations that make them the most likely to be involved in a murder, as a victim or an offender.[298]

immediately revealed that a very small number of neighborhoods in Chicago are responsible for most of the city's violence trends. The 'city's' crime problem is in fact geographically and socially concentrated in a few highly impoverished and socially isolated neighborhoods. Data also revealed that most victims (and offenders) of gun violence in Chicago tend to be young African American men who live in neighborhoods on the West or South sides of the city.").

[296] John Buntin, *Social Media Transforms the Way Chicago Fights Gang Violence*, GOVERNING, Oct. 2013, at 26, 28 ("Today, the Chicago Police Department is doing something similar with gangs. Using a tool academics call 'network analysis,' the CPD is mapping the relationships among Chicago's 14,000 most active gang members. It's also ranking how likely those people are to be involved in a homicide, either as victims or offenders. In the process, the CPD has discovered something striking: Cities don't so much have 'hot spots' as 'hot people.' That finding is transforming the way the police do business in Chicago and has significant implications for how other cities should be policed.").

[297] Michael Sierra-Arevalo, *How Targeted Deterrence Helps Police Reduce Gun Deaths*, SCHOLARS STRATEGY NETWORK (June 3, 2013, 1:11 PM), http://thesocietypages.org/ssn/2013/06/03/targeted-deterrence, *archived at* http://perma.cc/GZ65-U25X ("The perpetrators of gun violence are also concentrated in particular sectors of the population. In places like Boston, more than 50% of all murders and 70% of all shootings are committed by about one percent of youth aged 15 to 24."); *see also id.* ("Initiatives like The Boston Gun Project and Chicago's Project Safe Neighborhoods allow police to concentrate their efforts on gang-affiliated individuals with previous criminal records."). *See generally* OFFICE OF JUVENILE JUSTICE & DELINQUENCY PREVENTION, U.S. DEP'T OF JUSTICE, PROMISING STRATEGIES TO REDUCE GUN VIOLENCE 26-33 (1999), *available at* http://www.cops.usdoj.gov/html/cd_rom/solution_gang_crime/pubs/PromisingStrategiestoReduceGunViolence.pdf (discussing Boston's strategy to reduce gun violence by targeting specific groups and geographic areas).

[298] Monica Davey, *Chicago Tactics Put a Major Dent in Killing Trend*, N.Y. TIMES, June 11, 2013, at A1; *see also* Mark Guarnio, *Can Math Stop Murder?*, CHRISTIAN SCI. MONITOR (July 20, 2014), http://www.csmonitor.com/USA/2014/0720/Can-math-stop-murder-video, *archived at* http://perma.cc/G3TA-9SPT (discussing predictive policing techniques in Chicago including sending officers to the houses of suspected gang leaders).

Those four hundred individuals—part of a list of predicted offenders—were identified through big data techniques. Chicago police call it a "heat list."[299] Young men on the heat list are targets of predictive intervention-based strategies.

While a Fourth Amendment stop based solely on an individual's inclusion on this list, without more, might not be sufficiently particularized, big data tools exist to generate the necessary reasonable suspicion.[300] For example, imagine one of those four hundred individuals is a young man whom police wish to stop because they suspect that he is up to no good (and likely in possession of a gun). Plainly, an officer's suspicion that someone is "up to no good" does not constitute constitutionally sufficient justification for a stop. An officer sees the young man on the streets (but not engaged in any overt criminal activity). The officer identifies the young man as being on a list of individuals that predictive analytics suggested are at a heightened risk of involvement in gun violence. A quick NCIC database search reveals gang contacts, criminal associates, and prior arrests—including gun charges. Gang tattoos link the young man to local gangs. A license plate reader places the family car in the general vicinity of a gang shooting in the last month. His Facebook profile contains statements that police could interpret as directing violence at rival gang members.[301] Finally, predictive policing software has forecast the young man's location as the site of likely gun violence. If the police officer stops the young man (doing nothing overtly criminal) and finds a gun during a frisk, would a court really say there was not individualized and particularized suspicion that this individual was involved in gun and gang-related activity? Though the young man took no action to signify criminal activity, the data suggest that he was far more likely to be in possession of a gun than most people in Chicago.

How courts resolve these issues will determine the impact of big data on law enforcement. On one hand, judges might require some affirmative, imminent suspicious activity correlating with gun possession before upholding the stop, such as "furtive movements," a suspicious bulge, or unexplained

[299] Jeremy Gorner, *Chicago Police Use "Heat List" As Strategy to Prevent Violence*, CHI. TRIB. (Aug. 21, 2013), http://articles.chicagotribune.com/2013-08-21/news/ct-met-heat-list-20130821_1_chicago-police-commander-andrew-papachristos-heat-list, *archived at* http://perma.cc/8TJA-Y6KM.

[300] Presence on the list might also allow police to identify individuals for whom therapeutic intervention might be necessary.

[301] *See, e.g.*, JAAP BLOEM ET AL., SOGETI TREND LAB VINT, BIG SOCIAL: PREDICTING BEHAVIOR WITH BIG DATA 35 (2012), *available at* http://blog.vint.sogeti.com/wp-content/uploads/2012/10/Big-Social-Predicting-Behavior-with-Big-Data.pdf ("In the Netherlands, police officers go on duty with a smartphone in order to be able to pick up signals in the neighborhood from social media. In this way, they can show their faces before something serious happens in the schoolyard, for example.").

nervousness.[302] Without the requirement of some observable activity, the odds increase that predictive stops will target innocent people, criminalize by association, and negatively impact individuals based on little more than a hunch supported by non-criminal facts. On the other hand, many judges might find this totality of suspicions—even if focused on a particular suspect and not a crime—sufficient to justify an investigatory stop. Reasonable suspicion is a low threshold. Thus, in practice, aggregated reasonable suspicion would likely justify a stop in many courtrooms. As Part IV explains, this shift has significant implications for the Fourth Amendment.

D. *Big Data Suspicion*

Big data's ability to generate information about an identified suspect reveals the inherent vulnerability in the reasonable suspicion standard. Indeed, along the continuum of suspicion, more data makes it easier to satisfy the standard for two primary reasons. First, under a totality-of-circumstances test, the more factors a court considers in the totality, the easier it is to articulate suspicion. Quantity can make up for quality.[303] Second, the information provided by big data is individualized and particularized, consistent with the *Terry* language.[304] To be clear, the data are individualized to the criminal, not the crime. As courts apply *Terry*, however, which arose in the unknown suspect context, the difference becomes blurred.

This latter point is important to emphasize. The language in the earlier reasonable suspicion cases speaks to a general suspicion of unspecified criminal actions, using terms like "criminal activity may be afoot,"[305] "involved in criminal activity,"[306] "legal wrongdoing,"[307] or "illegality."[308] The general language does not require discussion of a particular observed crime (e.g., drug distribution or gun possession), because the officer actually observed the illegal activity in question. The observed crime and the

[302] *See, e.g.*, Jackson v. United States, 56 A.3d 1206, 1209-12 (D.C. 2013) (discussing the difficulty of interpreting furtive gestures and nervousness).

[303] Jane Bambauer, *Hassle*, 113 MICH. L. REV. (forthcoming 2014) (manuscript at 5), *available at* http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2404088 (recognizing "courts' consistent preference for police narratives chock full of detail, even when each additional detail does not contribute much to the suspicion"); *see also id.* (manuscript at 42) ("When assessing an officer's decision to stop or search somebody, courts prefer a long lists [sic] of reasons. The more reasons the agent can recount, the better.").

[304] Terry v. Ohio, 392 U.S. 1, 21 (1968).

[305] *Id.* at 30.

[306] Brown v. Texas, 443 U.S. 47, 51 (1979).

[307] United States v. Arvizu, 534 U.S. 266, 273 (2002).

[308] Florida v. J.L., 529 U.S. 266, 272 (2000).

observed criminal were not separate things to analyze; no distinction was needed in the analysis. Thus, in a small data world, the traditional language describing suspicious behavior has no meaning outside of the observed activity.

In a big data world, this same generalized language becomes distorted. An officer may know information about a suspect, but the question becomes: how does that information relate to the observed actions? Knowing someone is a "drug dealer" does not mean that the individual is actively dealing drugs at the moment of observation. Courts analyzing big data suspicion should thus be careful to require a direct link between the past data about a suspect and the observed suspicion. With big data suspicion, it is important for the individualized and particularized information to relate to the particular action observed. If a police officer identifies a suspect and learns information about the suspect's arrests, convictions, or associations that has nothing to do with the observed actions (if the officer observed any actions at all), then the new information should be irrelevant to the reasonable suspicion calculus. Only when those particularized factors can be connected to observed actions that signify criminal activity should they affect the analysis.

Courts will soon be asked to address the impact of big data on reasonable suspicion. But before that time, policymakers will need to think through and evaluate whether this innovation is good or bad for police, individuals, and society.

IV. Evaluating Big Data and Predictive Reasonable Suspicion

While big data may expose the fragility of the reasonable suspicion doctrine, the technology has arrived, and its impact on Fourth Amendment cases is imminent. As such, it is necessary to evaluate the questions of law and policy that arise from the move to big data policing. This Part discusses positives and negatives of big data policing and provides suggestions on how to address the pending evolution of Fourth Amendment doctrine.

A. *The Positives of Big Data Suspicion*

As may be evident from early adoption and experimentation with predictive technologies, law enforcement officials see the potential of these tools to reduce crime. Big data suspicion, if used correctly, can improve accuracy and efficiency, and it will yield unexpected insights into the patterns of criminal activity.

1. Improved Accuracy

Reasonable suspicion based on big data primarily benefits law enforcement because of the increased accuracy it purports to offer.[309] More information, and more precise information, should make it more likely that police target actual criminals rather than innocent people. In a small data environment, police rely on proxies for information to the detriment of everyone. Class, race, age, choice of clothing, and gender all factor into police officers' discretionary decisions on the street.[310] Police perceive ambiguous actions as suspicious because of subtle cues or instincts. These judgments also unfortunately include explicit and implicit biases, policing traditions, and the frailties of human perception.[311] Replacing those generalized intuitions with

[309] *Cf.* Rachael King, *IBM Analytics Help Memphis Cops Get "Smart,"* BLOOMBERG BUSINESSWEEK (Dec. 05, 2011), http://www.businessweek.com/technology/ibm-analytics-help-memphis-cops-get-smart-12052011.html, *archived at* http://perma.cc/Q77C-WCXW (describing the technology used by law enforcement in Memphis, Tennessee, which has contributed to the lowest crime rates there in a quarter-century).

[310] *See, e.g.*, Shima Baradaran, *Race, Prediction, and Discretion*, 81 GEO. WASH. L. REV. 157, 200 (2013) (examining "whether police demonstrate racial bias" in deciding whether to make arrests); Katherine Y. Barnes, *Assessing the Counterfactual: The Efficacy of Drug Interdiction Absent Racial Profiling*, 54 DUKE L.J. 1089, 1113, 1132-35 (2005) (explaining study results in which a driver's race was found to be "the most salient factor" in deciding whether to search a vehicle); Angela J. Davis, *Race, Cops, and Traffic Stops*, 51 U. MIAMI L. REV. 425, 425 (1997) (describing the reluctance of two black men to draw additional attention to themselves while driving because their race and gender already "makes them more likely to be stopped and detained by the police"); David A. Harris, Essay, *"Driving While Black" and All Other Traffic Offenses: The Supreme Court and Pretextual Traffic Stops*, 87 J. CRIM. L. & CRIMINOLOGY 544, 546, 570 (1997) ("[P]retextual police stops of blacks are so common—frequent enough to earn the name "driving while black"—[that] many African-Americans regularly modify the most casual aspects of their driving behavior . . . and even their personal appearance . . . ."); Noel Leader, Panel Discussion at CUNY School of Law (Sept. 29, 2010), *in Suspect Fits Description: Responses to Racial Profiling in New York City*, 14 CUNY L. REV. 57, 65-67 (2010) (asserting that illegal stops based on racial profiling are breaches of officers' duty, though police often attempt to justify them by citing alternative explanations like the suspect's dress); Tracey Maclin, Terry v. Ohio*'s Fourth Amendment Legacy: Black Men and Police Discretion*, 72 ST. JOHN'S L. REV. 1271, 1279-87 (1998) (arguing that although methods in place in the 1960s to deter crime were facially race-neutral, the implementation of these strategies was largely determined by the race of the subject).

[311] *See* Andrew E. Taslitz, *Police Are People Too: Cognitive Obstacles to, and Opportunities for, Police Getting the Individualized Suspicion Judgment Right*, 8 OHIO ST. J. CRIM. L. 7, 15-16 (2010) ("It is true that some people are, at times, reasonably good at making certain judgments based on first impressions. But they are also often quite bad at doing so. Moreover, first impressions can involve at least five major attributes, namely, the subject's emotions, personality, intelligence, mental states, and use of deception."); *id.* at 16 ("In addition, individuals' self-knowledge about the relative degree of accuracy of their ability to make judgments concerning each of the five major attributes upon first impression is also poor."); *see also* L. Song Richardson, *Police Efficiency and the Fourth Amendment*, 87 IND. L.J. 1143, 1147 (2012) ("Implicit social cognition research demonstrates that people have nonconscious reactions to others that can negatively influence their behaviors. These implicit biases begin when people categorize others both consciously and nonconsciously by

precise detail about actual people should result in a more accurate policing strategy. Humans are notoriously bad at making snap judgments, and while police officers make more snap judgments than most, they are not immune from the imperfections of human nature.[312]

For example, stories of racial profiling involving famous celebrities, wealthy professionals, and other citizens show how racial stereotypes can influence suspicion.[313] Yet in a big data world, a quick license plate scan or facial recognition check and a query of other databases (perhaps including professional licenses or even addresses), could help avoid the indignity of detention based solely on a police hunch.[314] Of course in many cases, information will not reveal that the individual is a celebrity, but even basic

race, gender, or a host of other socially relevant categories. Categorization triggers implicit stereotypes and attitudes." (footnotes omitted)).

[312] Eli B. Silverman, *With a Hunch and a Punch*, 4 J.L. ECON. & POL'Y 133, 140 (2007) ("Like other individuals within the same occupation, police vary in their ability to make intelligent, intuitive choices. Just as it varies among the general population, some police are better than others in detecting patterns from experience. Research and empirical observation amply demonstrates that there is a wide range in the ability of police officers to successfully deploy reasonable hunches in their work."); *see also* Anthony G. Greenwald & Linda Hamilton Krieger, *Implicit Bias: Scientific Foundations*, 94 CALIF. L. REV. 945, 947 (2006) (discussing the effects of mental processes outside of "conscious attentional focus" on decisionmaking); L. Song Richardson, *Cognitive Bias, Police Character, and the Fourth Amendment*, 44 ARIZ. ST. L.J. 267, 271 (2012) ("It is highly probable that fundamental attribution error affects police judgments of criminality. Officers on the beat often make quick decisions based upon limited evidence. The stressful nature of their jobs likely depletes their cognitive capacities, making correction for fundamental attribution error more difficult."); *id.* at 269 ("It is well established in the psychological literature that people tend to explain the behaviors of others by reference to their character (disposition) rather than to situational influences.").

[313] *See* CHARLES J. OGLETREE, JR., THE PRESUMPTION OF GUILT: THE ARREST OF HENRY LOUIS GATES JR. AND RACE, CLASS, AND CRIME IN AMERICA, 129-241 (2010) (telling the stories of one hundred influential African Americans who faced racial profiling or discrimination); David A. Harris, *The Stories, the Statistics, and the Law: Why "Driving While Black" Matters*, 84 MINN. L. REV. 265, 273-74 (1999) (describing measures taken by African Americans to avoid police harassment while driving); Sheri Lynn Johnson, *Race and the Decision To Detain a Suspect*, 93 YALE L.J. 214, 214 (1983) ("Thirty years ago police stopped Malcolm X because he was a black man in a white neighborhood. A revolution in civil rights later, police still view race as an important factor in the decision to detain a suspect." (footnote omitted)).

[314] *Compare* Albert W. Alschuler, *The Upside and Downside of Police Hunches and Expertise*, 4 J.L. ECON. & POL'Y 115, 118-19 (2007) (acknowledging that while hunches may be developed from real world experience, they are unreliable, shaped by racial stereotypes, burdensome to law enforcement, and unreviewable), *and* Harold Baer, Jr., *Got a Bad Feeling? Is That Enough? The Irrationality of Police Hunches*, 4 J.L. ECON. & POL'Y 91, 103 (2007) ("Until law enforcement agencies spend more time and money addressing the problems that arise from their culture, training and, in some locales, education, the hunch will remain problematical and occasionally unjust."), *with* Craig S. Lerner, *Judges Policing Hunches*, 4 J.L. ECON. & POL'Y 25, 25 (2007) ("[E]motions and intuitions are not obstacles to reason, but indispensable heuristic devices that allow people to process diffuse, complex information about their environment and make sense of the world.").

personal or employment data (or lack of criminal information) might provide police with a clue that a suspect is just an ordinary citizen not involved in criminal activity.

While vulnerable to abuse, predictive suspicion ultimately may make police stops more reliable. At its core, reasonable suspicion is a doctrine of predictive suspicion. The collected totality of circumstances must justify an officer's prediction that criminal activity is afoot.[315] Thus, having more information about an individual should result in more reliable predictions.[316] If police focus their efforts on people placed on "bad guy lists," it may protect individuals who are not on the lists. If police are forced to use big data to identify and link a suspect to a crime, they may also see patterns that suggest that the suspect was not involved in the crime. In this way, big data policing may be a measure more protective of individuals on the street.

The accuracy that big data provides not only increases the likelihood that police target the right suspects, but also, in turn, prevents the resulting physical, face-to-face interactions that generate tension.[317] Many police stops involve confirming or disproving suspicion.[318] Even if no arrest results, the unpleasant (and perhaps unnecessary) police–citizen contact breeds resentment and distrust.[319] Allowing police to confirm a person's lack

[315] *Cf.* Andrew E. Taslitz, *Fortune-Telling and the Fourth Amendment: Of Terrorism, Slippery Slopes, and Predicting the Future*, 58 RUTGERS L. REV. 195, 201 (2005) ("What is less often emphasized is that *Katz* faced the Justices with the question whether it is possible to authorize a search for non-existent evidence—evidence that may or may not come into being in the future.").

[316] *But cf.* Steinbock, *supra* note 18, at 38 ("The Fourth Amendment permits interferences with liberty and privacy based on predictions, often made by field officers, without notice to or consultation with the suspect.").

[317] *See* Frank Rudy Cooper, *"Who's the Man?": Masculinities Studies,* Terry *Stops, and Police Training*, 18 COLUM. J. GENDER & L. 671, 729-32 (2009) (criticizing police training programs for cultivating the culture of machismo and militarism that leads to police violence); James Forman, Jr., *Community Policing and Youth as Assets*, 95 J. CRIM. L. & CRIMINOLOGY 1, 14 (2004) (discussing police–citizen tension caused by "[b]elittling remarks, illegitimate orders, and cursing" by police during stops).

[318] *See* Minnesota v. Dickerson, 508 U.S. 366, 373 (1993) ("[W]here a police officer observes unusual conduct which leads him reasonably to conclude in light of his experience that criminal activity may be afoot . . . the officer may briefly stop the suspicious person and make reasonable inquiries aimed at confirming or dispelling his suspicions." (internal quotation marks omitted)).

[319] Mich. Dep't of State Police v. Sitz, 496 U.S. 444, 465 (1990) (Stevens, J., dissenting) ("[T]hose who have found—by reason of prejudice or misfortune—that encounters with the police may become adversarial or unpleasant without good cause will have grounds for worrying at any stop designed to elicit signs of suspicious behavior. Being stopped by the police is distressing even when it should not be terrifying, and what begins mildly may by happenstance turn severe."); David Rudovsky, *Law Enforcement by Stereotypes and Serendipity: Racial Profiling and Stops and Searches Without Cause*, 3 U. PA. J. CONST. L. 296, 334 (2001) ("[I]t is precisely at this intersection of crime, race and, police stop and frisk practices that the underlying social and legal conflicts most often are manifested, and not infrequently in sharp and violent confrontations.").

of involvement in criminal activity through a database search, rather than a physical stop, avoids unnecessary conflict.

2. Exculpatory Facts

Suspicion is not a one-way street. Suspicion can be disproved. Suspicion can be alleviated. The advent of big data suspicion may require consideration of exculpatory factors that lessen suspicion. Just as big data enables a wealth of suspicious inferences, it also generates an equal number of potentially exculpatory facts. For example, if Detective McFadden learned that John Terry's wife worked near the downtown location of the observation, pacing outside a store might turn from "casing a robbery" to "waiting for a loved one."

The potentially exculpatory nature of big data is a strong positive argument for its use in policing. Presumably, if big data information exists about a suspect, police should be obligated to check before initiating a stop.[320] The totality of circumstances should not be understood as the totality of suspicious activities; it should include exculpatory information that reduces suspicion as well. This is an established part of the probable cause analysis,[321] and big data technology allows it to be included in the reasonable suspicion analysis. Thus, existing exculpatory information should be factored into the totality of circumstances and weighted just as heavily as suspicious factors.

Courts might even require police to factor in exculpatory information as a self-contained check on the regular discretionary powers granted to police. When big data is available, an officer who did not use it in an exculpatory manner might be deemed to have acted recklessly.[322] In the same way that courts may take a negative inference from an unrecorded confession in a jurisdiction where videotaping confessions is the norm, a failure to use the available search technology might be held against the officer.[323] In this way,

[320] *See, e.g.*, Joshua A.T. Fairfield & Erik Luna, *Digital Innocence*, 99 CORNELL L. REV. 981, 1031 (2014) (arguing that defendants have a right to government-created exculpatory big data).

[321] *See, e.g.*, United States v. Grubbs, 547 U.S. 90, 95 n.2 (2006) ("[P]robable cause may cease to exist after a warrant is issued. The police may learn, for instance, that contraband is no longer located at the place to be searched."); United States v. Watson, 423 U.S. 411, 432 n.5 (1976) (Powell, J., concurring) ("But in some cases the original grounds supporting the warrant could be disproved by subsequent investigation that at the same time turns up wholly new evidence supporting probable cause on a different theory.").

[322] *See, e.g.*, Andrew Guthrie Ferguson, *Constitutional Culpability: Questioning the New Exclusionary Rules*, 66 FLA. L. REV. 623, 648-52 (2014) (discussing recklessness in the context of Fourth Amendment violations).

[323] *Cf., e.g.*, Steven A. Drizin & Beth A. Colgan, *Let the Cameras Roll: Mandatory Videotaping of Interrogations Is the Solution to Illinois' Problem of False Confessions*, 32 LOY. U. CHI. L.J. 337, 385-88

using big data to determine reasonable suspicion might actually prevent certain stops that would have been allowed under a traditional, small data reasonable suspicion standard.

3. Accountability

Focusing on big data sources also provides the potential for increased documentation of stops. In general, police do not document their suspicions *before a stop*, nor does anyone do so on their behalf.[324] Data-driven suspicion, though, can be documented beforehand. Police officers could demonstrate the steps they took to investigate a suspect by producing records of which databases they accessed and which search queries they used. In this way, police would replace the ex post justification for a stop with an ex ante description of the steps taken to validate a hunch before conducting a stop. In simple terms, a police officer could show that she checked the NCIC database and ran a license plate check before explaining why this information corroborated her initial suspicion. This record has the potential not only to limit whom police stop, but also to make a judge's determination of an officer's reasonable suspicion significantly easier.

This documentation will also encourage the development of a culture that allows for auditing of data, standards for record collection, and perhaps even notice requirements for targeted suspects. For example, police administrators, as an internal monitoring strategy, might examine an officer's history of stops to see what factors influenced his decision to stop a suspect. Looking through the documented history of big data searches and comparing them with the justifications for a stop might help police develop better training tools and build stronger accountability measures. Independent of any court case, internal monitoring measures can improve hit rates for arrests. In other data-driven contexts, these types of retention and accountability efforts are built into the regulating structure.[325]

(2001) (discussing a proposed law in Illinois that would have required videotaping confessions for certain crimes and made inadmissible confessions not videotaped).

[324] Many police officers are required to document certain police–citizen encounters after the fact. Jeffrey Fagan & Garth Davies, *Street Stops and Broken Windows:* Terry*, Race, and Disorder in New York City*, 28 FORDHAM URB. L.J. 457, 487-88 (2000) (describing the NYPD's use of UF-250 cards to record police–citizen encounters).

[325] *Cf., e.g.*, ABA STANDARDS FOR CRIMINAL JUSTICE: LAW ENFORCEMENT ACCESS TO THIRD PARTY RECORDS § 25-7.1 (3d ed. 2013), *available at* http://www.americanbar.org/content/dam/aba/publications/criminal_justice_standards/third_party_access.authcheckdam.pdf (recommending accountability mechanisms for databases used by law enforcement).

4. Efficiency

A move toward data-driven policing will also improve the efficient use of police resources. In many cases, better information will lead police to focus scarce resources on more serious risks and prevent unnecessary contacts with law-abiding citizens.

The rise of predictive policing signals the beginning of this shift to data-driven tips.[326] Underlying the theory of predictive policing is the idea that areas statistically more likely to have crime should have an additional police presence.[327] The data guide the officer patrol patterns, down to the particular time, date, and location.[328] While not replacing police patrols in other areas, police seek to target the areas identified through data.[329] Police administrators

[326] *See generally* Ferguson, *Predictive Policing*, *supra* note 29, at 265-69 (discussing the use of algorithms to predict crime and allocate law enforcement resources).

[327] *See* Braga et al., *supra* note 225, at 9 ("Criminological evidence on the spatial concentration of crime suggests that a small number of highly active micro places in cities—frequently called 'hot spots'—may be primarily responsible for overall citywide crime trends."); *see also* Joel M. Caplan et al., *Risk Terrain Modeling: Brokering Criminological Theory and GIS Methods for Crime Forecasting*, 28 JUST. Q. 360, 364 (2011) ("While a crime event occurs at a finite place, risk is a continuous dynamic value that increases or decreases intensity and clusters or dissipates in different places over time, even places remote from a crime event."); Shane D. Johnson et al., *Offender as Forager? A Direct Test of the Boost Account of Victimization*, 25 J. QUANTITATIVE CRIMINOLOGY 181, 184 (2009) (positing that the clustering of crimes could be explained by optimal foraging strategies); Shane D. Johnson et al., *Space–Time Patterns of Risk: A Cross National Assessment of Residential Burglary Victimization*, 23 J. QUANTITATIVE CRIMINOLOGY 201, 203-04 (2007) ("Most criminals commit crimes in areas with which they are already familiar."); Ashley B. Pitcher & Shane D. Johnson, *Exploring Theories of Victimization Using a Mathematical Model of Burglary*, 48 J. RES. CRIME & DELINQ. 83, 85-86 (2011) (discussing two theories that seek to explain the near-repeat phenomenon).

[328] *See generally* Goode, *supra* note 226, at A11 (reporting on anticipatory police deployments in Santa Cruz, California); *Predictive Policing: Don't Even Think About It*, ECONOMIST, July 20, 2013, at 24, 24-26 (describing data-driven police resource allocation); Leslie A. Gordon, *Predictive Policing May Help Bag Burglars—But It May Also Be a Constitutional Problem*, ABA JOURNAL (Sept. 1, 2013, 3:40 AM), http://www.abajournal.com/magazine/article/predictive_policing_may_help_bag_burglars--but_it_may_also_be_a_constitutio/, *archived at* http://perma.cc/J3L3-U9NN (discussing constitutional concerns relating to predictive policing); Kaste, *supra* note 219 (reporting on forward-looking policing strategies used in Seattle and other cities).

[329] *See* Charlie Beck & Colleen McCue, *Predictive Policing: What Can We Learn from Wal-Mart and Amazon About Fighting Crime in a Recession?*, POLICE CHIEF (Nov. 2009), http://www.policechiefmagazine.org/magazine/index.cfm?fuseaction=display_arch&article_id=1942&issue_id=112009, *archived at* http://perma.cc/D6DP-UYAS ("Predictive policing allows command staff and police managers to leverage advanced analytics in support of meaningful, information-based tactics, strategy, and policy decisions in the applied public safety environment. As the law enforcement community increasingly is asked to do more with less, predictive policing represents an opportunity to prevent crime and respond more effectively, while optimizing increasingly scarce or limited resources, including personnel.").

have embraced predictive policing because it allows them to allocate resources more efficiently while at the same time reducing crime.

Similarly, collecting data on individuals or groups perceived to be at high risk of entering the criminal justice system allows for a more focused use of police resources. Joint federal and state fusion centers have evolved to tackle gang and gun violence.[330] In these collaborative centers, police, with the help of technology, identify and map individuals by known gang associations, ethnicity, age, race, address, and social connections.[331] In Washington, D.C., one early partnership between federal and local law enforcement resulted in a "gang audit" that "helped identify 136 of the most violent gang/crew members in three of the highest crime areas in D.C."[332] People identified by police as involved in gangs faced targeted interventions, including face-to-face meetings, evictions from public housing, and criminal prosecution.[333] By mapping and targeting only those statistically most likely to be involved in criminal activity, the police attempted to address violence in the community proactively. This was the thinking behind the Chicago "heat list," and the approach has the benefit of focusing resources on those more likely to be involved in crime—whether as perpetrators or victims. While predictive policing practices raise a host of fairness concerns, from an efficiency perspective, recent innovations appear to have been a success.

5. Unexpected Insights

Big data also allows for unexpected insights from the collection of vast amounts of seemingly innocuous information. To package crack cocaine, a drug dealer needs tiny plastic bags and a scale.[334] To fire a gun, a shooter

[330] U.S. DEP'T OF JUSTICE, FUSION CENTER GUIDELINES: DEVELOPING AND SHARING INFORMATION AND INTELLIGENCE IN A NEW ERA F-3 (2006), *available at* http://www.it.ojp.gov/documents/fusion_center_guidelines.pdf (defining a fusion center as "[a] collaborative effort of two or more agencies that provide resources, expertise, and/or information . . . with the goal of maximizing the ability to detect, prevent, apprehend, and respond to criminal and terrorism activity"); *see also* Mimi Hall, *Feds Move to Share Intelligence Faster*, USA TODAY, July 27, 2006, at 3A (reporting that state fusion centers are run by "state police, FBI agents, National Guard, health officials and others").

[331] *Cf., e.g.*, Kelly, *supra* note 152 (describing the rise in use of cell phone information-gathering devices by police departments).

[332] SCOTT DECKER ET AL., PROJECT SAFE NEIGHBORHOODS: STRATEGIC INTERVENTIONS 18 (2007), *available at* https://www.bja.gov/publications/strategic_prob_solving.pdf.

[333] *Id.* at 17-19.

[334] *See* United States v. Dingle, 114 F.3d 307, 309 (D.C. Cir. 1997) (recounting expert testimony on practices used by drug dealers for packaging and distributing crack cocaine).

needs a bullet.[335] To break into a car, a thief needs tools (modern or old fashioned).[336] By tracking the sale of these items, police can recognize patterns and thus identify the criminals making the purchases. Similarly, most major criminal enterprises must launder money and otherwise hide illicit proceeds.[337] Unusual deposits, purchases, or money transfers can allow police to identify money laundering and the people involved.[338]

Incorporating geographic data can reveal patterns of location in an otherwise fluid criminal environment. Knowing where particular crimes occur can allow for more targeted suppression strategies. Big data allows for better tracking of national (or transnational) crimes, including human trafficking, drug smuggling, and credit card fraud.[339] For example, Google and others have partnered with three international antitrafficking nonprofits to track where calls for assistance originate to better map and disrupt human trafficking.[340] Similarly, hospital overdose admissions could reveal

[335] For an interesting story on how data about guns used in violent crime can be tracked and studied, see David S. Fallis, *Tracing Secrets*, WASH. POST, Oct. 24, 2010, at A1, which reports the findings of a *Washington Post* investigation into the sources of guns used in crimes—most notably that one dealer sold more than 2500 guns later recovered by police.

[336] *See, e.g.*, *Today: Rossen Reports* (NBC television broadcast June 5, 2013), *available at* http://www.today.com/news/police-admit-theyre-stumped-mystery-car-thefts-6C10169993 (reporting on a series of car thefts committed using a device that quickly bypasses electronic locks).

[337] Jimmy Gurulé, *The Money Laundering Control Act of 1986: Creating a New Federal Offense or Merely Affording Federal Prosecutors an Alternative Means of Punishing Specified Unlawful Activity?*, 32 AM. CRIM. L. REV. 823, 823 (1995) (describing money laundering as the "lifeblood" of organized crime); *see also Money Laundering Legislation: Hearing on S. 572, S. 1335, and S. 1385 Before the S. Comm. on the Judiciary*, 99th Cong. 30 (1985) (statement of Sen. DeConcini, Member, S. Comm. on the Judiciary) ("Without the means to launder money, thereby making cash generated by a criminal enterprise appear to come from a legitimate source, organized crime could not flourish as it now does.").

[338] *See, e.g.*, Richard K. Gordon, *Losing the War Against Dirty Money: Rethinking Global Standards on Preventing Money Laundering and Terrorism Financing*, 21 DUKE J. COMP. & INT'L L. 503, 527-28 (2011) (describing the "red flags" used by the Treasury Department's financial intelligence unit, FinCEN, to identify money laundering).

[339] *See, e.g.*, Philip K. Chan et al., *Distributed Data Mining in Credit Card Fraud Detection*, IEEE INTELLIGENT SYSTEMS, Nov.–Dec. 1999, at 67, 68 (providing technical details of specific credit card fraud identification algorithms); Scott R. Peppet, *Prostitution 3.0?*, 98 IOWA L. REV. 1989, 2039-40 (2013) (suggesting data with which to estimate the likelihood that a prostitute is a victim of human trafficking); Peter P. Swire, *Privacy and Information Sharing in the War on Terrorism*, 51 VILL. L. REV. 951, 964 (2006) (discussing the "out of pattern" system for identifying credit card fraud).

[340] Bernhard Warner, *Google Turns to Big Data to Unmask Human Traffickers*, BLOOMBERG BUSINESSWEEK (Apr. 10, 2013), http://www.businessweek.com/articles/2013-04-10/google-turns-to-big-data-to-unmask-human-traffickers, *archived at* http://perma.cc/3CSC-RDUJ ("The [Google-led] alliance . . . means the three anti-trafficking networks . . . will share data on where the emergency phone calls are originating, the ages of the victims, their home countries, and the types of criminal activities they have been forced into. . . . [T]he agencies will be able to crunch data

drug use patterns. Social media trends may reveal clues about gang activities,[341] prostitution services,[342] or cybercrime.[343]

Patterns of crime can also reveal the locations of criminal actors. Police can link certain getaway routes to robbery hotspots.[344] Locations of gunshots can reveal shifting gang-related turf borders.[345] Social services visits to monitor "stay-away orders" can reveal potential locations of future domestic violence.[346] Even the type of alcohol sold at stores and restaurants can correlate with the rate of violent crime in a neighborhood.[347] These insights can help police investigate and prevent crime and would not have been easily observed before the advent of big data.

like this in real time to detect crime trends that they can then share with police and policymakers to help protect victims.").

[341] *See, e.g.*, Ben Austen, *Public Enemies: Social Media Is Fueling Gang Wars in Chicago*, WIRED (Sept. 17, 2013, 6:30 AM), http://www.wired.com/2013/09/gangs-of-social-media/, *archived at* http://perma.cc/3L5H-L2M2 (describing escalating gang tensions via Twitter and YouTube).

[342] *See, e.g.,* Erica Fink & Laurie Segall, *Pimps Hit Social Networks to Recruit Underage Sex Workers*, CNNMONEY, (Feb. 27, 2013, 7:30 AM), http://money.cnn.com/2013/02/27/technology/social/pimps-social-networks, *archived at* http://perma.cc/S4BU-LEUK (reporting on the use of Facebook and other social media sites to lure victims into becoming sex workers).

[343] *See generally Online Privacy, Social Networking, and Crime Victimization: Hearing Before the Subcomm. on Crime, Terrorism, & Homeland Sec. of the H. Comm. on the Judiciary*, 111th Cong. 5-12 (2010) (statement of Gordon M. Snow, Asst. Dir., FBI) (discussing ways in which cybercriminals use social media to deceive victims).

[344] JENNIFER BACHNER, IBM CTR. FOR THE BUS. OF GOV'T, PREDICTIVE POLICING: PREVENTING CRIME WITH DATA AND ANALYTICS 15-16 (2013), *available at* http://www.businessofgovernment.org/sites/default/files/Predictive%20Policing.pdf. (suggesting that criminals prefer "areas with desirable escape routes," including "[a]reas in close proximity to features such as interstate highways, bridges, and tunnels").

[345] *See* Andras Petho, David S. Fallis & Dan Keating, *Acoustic Sensors Reveal Hidden Depth of Gun Use in D.C.*, WASH. POST, Nov. 2, 2013, at A1 (describing data from the District of Columbia's acoustic "ShotSpotter" system, which had identified 39,000 separate instances of gunfire, many of which were clustered geographically).

[346] Goldstein, *supra* note 294, at A1 (discussing efforts by the NYPD to reduce domestic violence).

[347] Robert Lipton et al., *The Geography of Violence, Alcohol Outlets, and Drug Arrests in Boston*, 108 AM. J. PUB. HEALTH 657, 661 (2013) (suggesting "a positive relationship between violent crime and the presence of package stores," but "a negative relationship between violent crime and the presence of restaurants selling beer and wine."); *see also* Press Release, Univ. of Mich. Health Sys., Could a Computer on the Police Beat Prevent Violence? (Feb. 18, 2013), *available at* http://www.uofmhealth.org/news/archive/201302/could-computer-police-beat-prevent-violence ("Results from the study indicate that types and densities of alcohol outlets were directly related to violent crimes despite the fact that alcohol outlets are typically viewed as locations in which other population or environmental factors, such as poverty or prostitution, relate to the violence.").

B. *The Negatives of Big Data Suspicion*

While big data offers much promise, big data–driven policing also has potential negative consequences. This Section outlines a few representative concerns.

1. Bad Data

A system based on data requires accurate, up-to-date information.[348] One concern with a vast, ever-growing, networked data system is that the quality controls on shared data are almost nonexistent.[349] Police may rely on existing data without any knowledge of how the data was collected or whether mechanisms exist to ensure its accuracy. Data problems have emerged even within locally controlled systems[350] and certainly arise when jurisdictions share information.[351] Reputed "gang lists" used by police have been shown to be inaccurate.[352] Arrest reports can be inaccurate or erroneous

[348] Cope, *supra* note 24, at 193 ("Data quality affected the development of analysis. Analysts frequently found crucial details missing from intelligence reports for their products.").

[349] *See, e.g.*, Eric J. Mitnick, *Procedural Due Process and Reputational Harm: Liberty as Self-Invention*, 43 U.C. DAVIS L. REV. 79, 126 (2009) (noting that while most databases are supposed to be subject to quality control, "[i]n reality . . . , the evidence is overwhelming that the control measures currently in place regularly fail, either due to lack of resources, skill, or because they are simply neglected"); Wright, *supra* note 121, at 122 (finding quality control lacking in one database where no reports were questioned by superiors; the officers making some of the reports had no gang experience, and there were no reviews for accuracy).

[350] *See* Jeff Morganteen, *What the CompStat Audit Reveals About the NYPD*, N.Y. WORLD (July 3, 2013), http://www.thenewyorkworld.com/2013/07/03/compstat/, *archived at* http://perma.cc/K4ZP-KR4L ("The outside audit . . . not only confirmed that such data manipulation takes place but found several weak points in the ways the department tracks and uncovers it."); *see also* DAVID N. KELLEY & SHARON L. MCCARTHY, THE REPORT OF THE CRIME REPORTING REVIEW COMMITTEE TO COMMISSIONER RAYMOND W. KELLY CONCERNING COMPSTAT AUDITING 47 (2013), *available at* http://www.nyc.gov/html/nypd/downloads/pdf/public_information/crime_reporting_review_committee_final_report_2013.pdf ("[T]he patterns of the misclassified reports support in some measure the anecdotal accounts . . . that certain types of incidents may be downgraded as a matter of practice in *some* precincts.").

[351] Herring v. United States, 555 U.S. 135, 155 (2009) (Ginsburg, J., dissenting) ("The risk of error stemming from these databases is not slim. Herring's *amici* warn that law enforcement databases are insufficiently monitored and often out of date. Government reports describe, for example, flaws in NCIC databases, terrorist watchlist databases, and databases associated with the Federal Government's employment eligibility verification system." (footnotes and citation omitted)).

[352] *See, e.g.*, Mitnick, *supra* note 349, at 126; Wright, *supra* note 121, at 129 ("In sum, gang databases appear to be riddled with factual inaccuracies, administrative errors, lack of compliance with departmental guidelines, and lack of oversight.").

but remain in public and private databases.[353] The FBI's own files—used for millions of background checks—reportedly contain hundreds of thousands of errors.[354] Worse, there is no simple mechanism to clear the bad data from a web of networked systems all sharing the same errors.[355]

Adding private, third-party sources of information only compounds the problem. CBS News's 60 Minutes reported that "as many as forty million Americans have a mistake on their credit report. Twenty million have significant mistakes."[356] These are the same credit report datasets that underlie many commercial big data systems. Both discovering and correcting mistakes is difficult; it requires knowledge of the error and the wherewithal to change it. Police agents accessing records, however, would have no knowledge that an error existed—or even necessarily a way to check the accuracy of the data. Mistakes can occur at any point in the process from collection to entry to analysis. In addition, data can grow stale. Typographical errors can lead to erroneous linkages.[357] These mistakes can have real consequences on individual liberty. As Justice Ginsburg warned:

> Inaccuracies 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.[358]

[353] Roberto Concepción, Jr., *Need Not Apply: The Racial Disparate Impact of Pre-Employment Criminal Background Checks*, 19 GEO. J. ON POVERTY L. & POL'Y 231, 246-48 (2012) (highlighting the high cost of false positives in pre-employment queries of criminal records databases).

[354] *See* Ylan Q. Mui, *Use of FBI Database in Hiring Raises Concerns*, WASH. POST, July 30, 2013, at A1 (discussing a report by the National Employment Law Project on errors in FBI background checks).

[355] *See, e.g.*, Anita Ramasastry, *Lost in Translation? Data Mining, National Security and the "Adverse Inference" Problem*, 22 SANTA CLARA COMPUTER & HIGH TECH. L.J. 757, 775-76 (2006) (discussing reports of errors and inaccuracies in credit reports); Tal Z. Zarsky, *Governmental Data Mining and Its Alternatives*, 116 PENN ST. L. REV. 285, 298 (2011) (discussing the problem of errors in data mining processes).

[356] *60 Minutes: 40 Million Mistakes: Is Your Credit Report Accurate?* (CBS television broadcast Feb. 10, 2013), *available at* http://www.cbsnews.com/8301-18560_162-57567957/credit/.

[357] *Cf.* Wayne J. Pitts, *From the Benches and Trenches: Dealing with Outstanding Warrants for Deceased Individuals: A Research Brief*, 30 JUST. SYS. J. 219, 220 (2009) (describing a study that discovered numerous errors in a warrant database, including incorrect social security numbers, inaccurate names, and "illogical birth dates," and noting that "none of the[] issues are surprising or unusual given the nature of the population being tracked").

[358] Herring v. United States, 555 U.S. 135, 155-56 (2009) (Ginsburg, J., dissenting) (internal quotation marks omitted).

The lack of transparency in these data systems only increases the chance of error. Police systems are usually restricted to authorized police users. Private companies, seeking commercial gain, have little incentive to reveal the workings of proprietary systems or the data thereby collected. No agency has the responsibility to audit the growing governmental and commercial big data network. While the Federal Trade Commission has promised to monitor private big data companies,[359] it has little ability to examine the data itself and has no role in oversight of law enforcement use of the data. Though oversight institutions do exist (including courts, congressional committees, and independent agencies),[360] the volume of information at issue prevents these groups from examining the quality of the data or the magnitude of the errors.[361] Without transparency, there can be little hope for accountability to ensure that data systems will be sufficiently reliable to justify altering constitutional rights.[362] In short, big data suspicion may be based on bad data.

[359] *See* Brendan Sasso, *FTC Chief Targets Firms with Vast Databases*, HILL (Aug. 19, 2013, 9:12 PM), http://thehill.com/policy/technology/317729-ftc-chief-targets-firms-with-vast-databases, *archived at* http://perma.cc/8HTB-SE4W (reporting that the head of the FTC stated that the agency "will use its power to punish deceptive business practices [and] to crack down on firms that fail to live up to their own promises about how they will use their customers' data"). *See generally* FTC, DATA BROKERS: A CALL FOR TRANSPARENCY AND ACCOUNTABILITY 46-56 (2014), *available at* http://www.ftc.gov/system/files/documents/reports/data-brokers-call-transparency-accountability-report-federal-trade-commission-may-2014/140527databrokerreport.pdf (presenting findings of an FTC study of large data brokers and recommending reforms).

[360] For example, the House Select Committee on Intelligence and the Senate Select Committee on Intelligence have legislative oversight of the intelligence agencies. The House Committee on the Judiciary, the Senate Committee on the Judiciary, the House Committee on Homeland Security, the Senate Committee on Homeland Security and Governmental Affairs, and others have oversight of domestic surveillance. Independent agencies such as the Privacy and Civil Liberties Oversight Board have general oversight. The Foreign Intelligence Surveillance Court provides some judicial oversight. General counsels and inspectors general add additional layers of protection.

[361] For example, the Foreign Intelligence Surveillance Court released a redacted opinion offering insight into problems with overcollection of phone records by the National Security Agency. *See* [Redacted], [Redacted], 2011 WL 10945618 (FISA Ct. Oct. 3, 2011). In its October 2011 opinion, the court revealed that it could review only samples of the NSA-collected data due to the incredible number of search queries and volume of data involved with the NSA's operations. *See id.* at *10; *see also In re* Order Requiring Production of Tangible Things From [Redacted], 2013 WL 5741573, at *10-14 (FISA Ct. Aug. 29, 2013) (No. BR 13-109) (setting guidelines for review of NSA metadata-related surveillance programs); MAJORITY STAFF, SENATE COMM. ON COMMERCE, SCI. & TRANSP., 113TH CONGRESS., A REVIEW OF THE DATA BROKER INDUSTRY: COLLECTION, USE, AND SALE OF CONSUMER DATA FOR MARKETING PURPOSES (2013), *available at* http://www.commerce.senate.gov/public/?a=Files.Serve&File_id=0d2b3642-6221-4888-a631-08f2f255b577 (providing an example of a congressional investigation into data brokers and the collection of personal information).

[362] *See* Solove, *Data Mining*, *supra* note 177, at 359 ("Another key issue regarding the liberty side of the balance is transparency—the degree of openness by which a particular security measure

2. False Positives

Even assuming "good data," big data reasonable suspicion will result in false positives whereby police stop innocent people.[363] This burden will fall most heavily on individuals who have some criminal history, but who are not currently engaging in criminal activity.[364] Predictive analytics will suggest suspicion based on an identified correlation, but such suspicion will often be unfounded. Perhaps a license plate reader will place the car of a convicted burglar within a predicted burglary hotspot, which also happens to be next to the convicted burglar's grandmother's house. Police might stop the suspected burglar solely because of this correlation. One can imagine that those individuals who find themselves on a "bad guy list" will be marked for more than their fair share of borderline suspicious stops.[365]

Big data suspicion creates the real concern that certain individuals by virtue of their past criminal activities will always be at risk to be stopped. Those with lengthy criminal records or gang associations may be stopped because of who they are and not what they are doing. Prior police contacts will become the digital "scarlet letter" marking certain people in a community as suspicious.[366]

Over the past several decades, poor people and people of color have had disproportionate contact with the criminal justice system.[367] If these contacts become data points that can be used in a reasonable suspicion analysis, then these data may become proxies for race or class (with similar

is carried out. Transparency is essential to promote accountability and to provide the public with a way to ensure that government officials are not engaging in abuse.").

[363] *Cf.* Taslitz, *supra* note 311, at 10 ("Any concept of reasonable suspicion . . . that tolerates massive false negative rates—frequent invasions of privacy, property, and locomotive rights that ensnare the apparently innocent—is a flawed conception. The costs imposed on communities and individuals become great, while little in the way of crime-control efforts is achieved.").

[364] *See supra* Part III.

[365] Of course, these individuals might also be targeted without a big data–inspired list.

[366] *See* David Wolitz, *The Stigma of Conviction: Coram Nobis, Civil Disabilities, and the Right to Clear One's Name*, 2009 BYU L. REV. 1277, 1316 (arguing that a criminal conviction is a "uniquely stigmatizing piece of information" and that it disproportionately affects a person's reputational profile).

[367] *See* Robin Walker Sterling, *Raising Race*, CHAMPION, Apr. 2011, at 24, 24-25 ("The criminal justice system has exploded outside of the prison walls, as well. As of 2009, the number of people under criminal justice supervision—including those who are in jail, in prison, on probation, and on parole—totaled 7.2 million people. In a dismaying parallel to incarceration rates, people of color are also overrepresented among arrestees, probationers, and parolees. There are more African Americans under correctional control today than were enslaved in 1850. . . . With numbers like these, it is clear that this overrepresentation of minorities in the criminal justice system, or disproportionate minority contact (DMC), is one of the major human rights violations of our time." (footnotes omitted)).

effect). For example, the ACLU's recent national study on marijuana arrests demonstrates that African Americans are more likely to be arrested for marijuana than whites, despite equivalent usage rates.[368] Thus, more data has been collected about minority marijuana possession, even though whites commit the crime at the same rate. If data are collected only about certain classes of people, then those people are more likely to become future targets of suspicion simply because of the initial selection bias. Thus, important questions remain about who collects, interprets, and chooses the big data to study.[369]

Worse, like other quantitative systems used for decisionmaking, big data–based predictive policing will appear to be objective and fair when it may in fact reflect subjective factors and structural inequalities. Just as we have credit ratings that allow lenders to predict future creditworthiness, police could develop "criminal ratings" to predict future criminal proclivity.[370]

Similarly, data can lead us to believe our own worst instincts.[371] If published data demonstrate a higher arrest rate for people of color, then this information may well influence discretionary decisions about who to stop.[372] Implicit bias and confirmation bias will result in police seeing what they have been told to see, even if it is not actually occurring.[373] Implicit bias involves unconscious prejudices that influence individuals making discretionary

[368] ACLU, THE WAR ON MARIJUANA IN BLACK AND WHITE 17-22 (2013), *available at* https://www.aclu.org/files/assets/aclu-thewaronmarijuana-rel2.pdf (reporting that blacks are roughly four times more likely to be arrested for marijuana possession than whites despite similar usage rates); *see also* Steven Nelson, *ACLU Marijuana Study: Blacks More Likely to be Busted*, U.S. NEWS & WORLD REP. (June 4, 2013, 5:26 PM), http://www.usnews.com/news/newsgram/articles/2013/06/04/aclu-marijuana-study-blacks-more-likely-to-be-busted, *archived at* http://perma.cc/W5DS-ZBUG (reporting on the ACLU study).

[369] *See* Cohen, *supra* note 7, at 1922 ("It is beyond serious question that the techniques that comprise Big Data offer vitally important strategies for promoting human flourishing in an increasingly complex, crowded, and interdependent world. But those techniques cannot themselves decide which questions to investigate, cannot instruct us how to place data flows and patterns in larger conceptual or normative perspective, and cannot tell us whether and when it might be fair and just to limit data processing in the service of other values.").

[370] Thank you to the discussants at Northeastern University School of Law's Legal Scholarship 4.0 conference for developing the concept of "criminal ratings."

[371] Thank you to the discussants at the criminal law professor workshop at the Washington College of Law, American University, for developing this argument.

[372] *Cf.* Taslitz, *supra* note 311, at 44-45 (discussing the potential for extrapolation from past experience despite insufficient information).

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

decisions.[374] This can result in unequal outcomes for similarly situated individuals.[375] Implicit bias inevitably exists in the ordinary course of police activities, but is even more damaging when combined with confirmation bias: "the tendency to bolster a hypothesis by seeking consistent evidence while minimizing inconsistent evidence."[376] Thus, an officer conditioned to believe that a particular type of person may be more likely to commit a criminal act will likely see that person through the lens of suspicion. By providing the information to confirm this suspicion, big data will make it easier for police to justify a stop. Even more dangerously, an officer with discriminatory animus may be able to justify a knowingly unconstitutional stop using an aggregation of otherwise innocent data.[377]

This risk demonstrates how suspicions about past criminal actions can all too easily morph into suspicions about current criminal activity. It highlights the importance of requiring a nexus between the suspected criminal and the suspected criminal activity. It also highlights the dangers of how big data can target certain populations based on correlations with possible criminal activity, rather than causation from real criminal activity. Justification to stop these individuals—marked by big data—will be too easily met, undermining the individualized and particularized protections in the Fourth Amendment.

3. Shifting Power Balance

The Constitution establishes a power-sharing relationship between citizens and the government. The Fourth Amendment, like other parts of the Bill of Rights, represents a check on government power.[378] The probable cause standard, and to a lesser extent, the reasonable suspicion standard, limits the actions of government agents. Big data, by weakening the reasonable

[374] Richardson, *supra* note 312, at 271-72.

[375] *See* Mary Fan, *Street Diversion and Decarceration*, 50 AM. CRIM. L. REV. 165, 192 (2013) ("A rich body of literature has documented how implicit biases—negative perceptions of minorities that may unconsciously lurk despite best intentions—impact the judgment of an array of actors, such as police, prosecutors, and jurors.").

[376] Barbara O'Brien, *Prime Suspect: An Examination of Factors That Aggravate and Counteract Confirmation Bias in Criminal Investigations*, 15 PSYCHOL. PUB. POL'Y & L. 315, 315 (2009); *see also id.* at 318 (noting that "[p]olice investigators are also prone to confirmation bias").

[377] *See* Richard Winton et al., *LAPD to Build Data on Muslim Areas*, L.A. TIMES, Nov. 9, 2007, at A1 (describing a police initiative to identify areas "at-risk" for terrorist activities based on ethnicity); Richard Winton et al., *Outcry Over Muslim Mapping*, L.A. TIMES, Nov. 10, 2007, at A1 (same).

[378] *See* United States v. Martinez-Fuerte, 428 U.S. 543, 554 (1976) (noting that the purpose of the Fourth Amendment is to protect against "arbitrary and oppressive interference by enforcement officials with the privacy and personal security of individuals").