Saint Louis University Law Journal

Volume 65
Number 4 *Tradeoffs: Technology, Privacy, and the Law (Summer 2021)*

Article 5

2021

Dye in the Cracks: The Limits of Legal Frameworks Governing Police Use of Big Data

Sarah Brayne
The University of Texas at Austin, sbrayne@utexas.edu

Recommended Citation

Sarah Brayne, *Dye in the Cracks: The Limits of Legal Frameworks Governing Police Use of Big Data*, 65 St. Louis U. L.J. (2021).
Available at: https://scholarship.law.slu.edu/lj/vol65/iss4/5

DYE IN THE CRACKS: THE LIMITS OF LEGAL FRAMEWORKS GOVERNING POLICE USE OF BIG DATA[†]

SARAH BRAYNE*

Computational procedures increasingly inform how we work, communicate, and make decisions, raising sociolegal questions about how data are used by police and what the consequences are for laws governing police activity. Legal scholars have begun analyzing the implications of big data policing, yet their work to date is largely theoretical. In this article, I draw on ethnographic fieldwork conducted with the Los Angeles Police Department ("LAPD") to ground legal debates about police use of big data in empirical detail. The article opens with a brief description of the fieldwork and findings on how the police use big data for dragnet and directed surveillance. I then identify four ways legal frameworks are overlooking the social side of big data. First, the way the conceptual categories that underpin legal doctrine—like individualized suspicion—are deployed and organized to make normative assessments do not reflect how decision-making plays out on the ground. Second, police are not simply scaling up data collection in the digital age; rather, different kinds of data are being produced. Despite the fact that there is a difference in kind—rather than just degree—old legal doctrine is still being laid on top of these data. Third, relying on extant legal mechanisms like the exclusionary rule means using what is meant to be a check on state power at one point in time and space, whereas data is fundamentally social and, as such, has a life course. Fourth, unfettered big data policing creates new opportunities for information asymmetries and can threaten due process through parallel construction.

[†] Portions of this material were originally published in chapters 4 and 7 of *Predict and Surveil: Data, Discretion and the Future of Policing* and have been reproduced by permission of Oxford University Press. For more information, please visit https://global.oup.com/academic/product/predict-and-surveil-9780190684099 [https://perma.cc/F5CS-VHAX].

* University of Texas at Austin, Department of Sociology.

TABLE OF CONTENTS

INTRODUCTION

We are currently experiencing the intersection of two structural shifts—the growth of criminal justice surveillance and the rise of so-called "big data." As many have documented, the scope of the American criminal legal system—from policing to incarceration—has expanded rapidly over the past half century. Its reach is now unprecedented, both historically and in an international comparative perspective.[1] Approximately one third of the United States' adult population—over 70 million Americans—has a record on file with criminal justice agencies.[2]

In the past decade, we have also seen an explosion of interest in big data, or the computational analysis of massive and diverse datasets to automate decisions and make predictions. From finance to politics, credit, insurance, and medicine, actors across a range of institutional domains have started leveraging the vast troves of digital data we leave in our wake as we go about our everyday lives. Law enforcement is no exception, raising sociolegal questions about how data are used by police to surveil people, whether and how their use shifts discretion and organizational practice, and what the consequences of these shifts are for social inequality and the law.

Legal scholars have begun analyzing the legal implications of big data policing, yet their work to date is largely theoretical. Few law professors undertake fieldwork, conducting empirical research and watching how the police interpret and enact law in their everyday work.[3] In this article, I draw on ethnographic fieldwork conducted with the Los Angeles Police Department ("LAPD") in order to ground legal debates about police use of big data in empirical detail.

I argue that basic legal principles are inadequate for governing policing in the digital age not simply because they are anachronistic, but also because the legal debates are too narrow. By not attending to the sociological processes that underpin basic legal principles, they take existing legal categories for granted.

[1] E. Ann Carson, *Prisoners in 2014*, U.S. DEP'T OF JUST. BUREAU OF JUSTICE STATISTICS (Sept. 2015), https://www.bjs.gov/content/pub/pdf/p14.pdf [https://perma.cc/4CRW-ZJVJ]; DAVID GARLAND, THE CULTURE OF CONTROL: CRIME AND SOCIAL ORDER IN CONTEMPORARY SOCIETY 2 (2002); Sara Wakefield & Christopher Uggen, *Incarceration and Stratification*, 36 ANN. REV. OF SOCIO. 387, 388 (2010); NATIONAL RESEARCH COUNCIL, THE GROWTH OF INCARCERATION IN THE UNITED STATES: EXPLORING CAUSES AND CONSEQUENCES (2014); Erinn J. Herberman & Thomas P. Bonczar, *Probation and Parole in the United States, 2013*, U.S. DEP'T OF JUST.: BUREAU OF JUST. STATISTICS (2014), https://www.bjs.gov/index.cfm?ty=pbdetail&iid=5135 [https://perma.cc/4NXJ-UGXG].

[2] Gary Fields & John R. Emshwiller, *As Arrest Records Rise, Americans Find Consequences Can Last a Lifetime*, WALL ST. J. (Aug. 18, 2014, 10:30 PM), https://www.wsj.com/articles/as-arrest-records-rise-americans-find-consequences-can-last-a-lifetime-1408415402 [https://perma.cc/4DML-L4TX].

[3] For an exception, see Kathryne M. Young & Christin Munsch, *Fact and Fiction in Constitutional Criminal Procedure*, 66 S.C. L. REV. 445 (2014).

Put differently, big data—and the legal implications of its use—are not being considered in a sufficiently *sociological* way. As a result, as social life and its attendant data evolve, legal scholars are developing theories to explain the growing gulf between the way the modern courts understand data and the central concerns of foundational legal concepts such as reasonable suspicion, the Fourth Amendment, and the third-party doctrine. However, these theories are insufficiently sociological. In addition to considering how legal constructs *should* adapt, we should ask whether constitutional modes of police regulation are constructed in such a way that they *can* adapt.

In this article, I identify four ways legal frameworks are overlooking the social side of big data. First, the way the conceptual categories that underpin legal doctrine—like individualized suspicion—are deployed and organized to make normative assessments do not reflect how decision-making plays out on the ground. Second, police are not simply scaling up data collection in the digital age; rather, different kinds of data are being produced. Despite the fact that there is a difference in kind—rather than just degree—old legal doctrine is still being laid on top of these data. Third, relying on extant legal mechanisms like the exclusionary rule means using what is meant to be a check on state power at one point in time and space, whereas data is fundamentally social and, as such, has a life course. Fourth, unfettered big data policing creates new opportunities for information asymmetries and can threaten due process through parallel construction.

I. FIELDWORK

Ironically, there is little data on police use of big data. Therefore, to understand how the law enforcement is using big data in their daily operations, I conducted ethnographic fieldwork with the LAPD, an agency on the leading edge of police use of advanced analytics. Between 2013 and 2018, I conducted observations and interviews with sworn officers and civilian employees in area divisions and specialized divisions including robbery-homicide, information technology, fugitive warrants, records and identification, juvenile, and risk management divisions, as well as at the Real Time Crime Analysis Center ("RACR"), the Emergency Operations Center in Downtown Los Angeles. I shadowed analysts as they worked with data, interviewed officers, and went on ride-alongs in patrol cars and a helicopter to see how officers used data in the field. I also went to the Joint Regional Intelligence Center ("JRIC"), the fusion center in Southern California, interviewed individuals in LA County Sherriff's Department, and individuals who work at technology firms that design key analytic software the LAPD uses, such as Palantir and PredPol.

II. HOW DO THE POLICE USE BIG DATA?

The LAPD uses big data for two types of surveillance, *dragnet* and *directed surveillance*. Dragnet surveillance refers to surveillance tools that gather information on everyone, rather than just those under suspicion. An example of a dragnet surveillance tool is the automatic license plate reader ("ALPR") which takes two photos of every vehicle that passes through its line of vision—one of the car and one of the license plate—and records the time, date, and coordinates. Police can compare ALPR readings against lists of outstanding warrants or stolen cars, or use them in the course of investigations to make inferences about individuals' locations, travel patterns, or temporary residences.

Directed surveillance, by contrast, refers to the surveillance of people and places deemed suspicious. An emblematic example of directed surveillance is predictive policing. The LAPD conducted both place- and person-based predictive policing. The department used a place-based predictive policing algorithm designed by PredPol to predict property crime. The algorithm uses three different inputs—past place, type, and time of crime—as training data, weighing more recent crimes more heavily than older ones, in order to produce 500 square foot boxes where future property crime is more likely to occur. Officers are given PredPol maps at the beginning of their shifts and directed to drive to predictive boxes during their uncommitted time (i.e., when they are not responding to a call for service or booking someone).

Whereas place-based predictive policing was primarily used for property crime, person-based predictive policing was primarily used for violent crime. In the LAPD, they used a federally funded point system called Operation LASER. The Crime Intelligence Detail ("CID") first plots crimes in their division and identifies a "problem" crime. Then, they shift their unit of analysis from crimes to individuals, gathering intelligence daily from patrols, the Parole Compliance Unit, field interview cards (police contact cards), traffic citations, release from custody forms, crime and arrest reports, and criminal histories. These data are used to generate a list of "chronic offenders." Each person on the list is assigned a point value and rank ordered according to that value. Individuals are assigned five points for a violent criminal history, five points for known gang affiliation, five points for prior arrests with a handgun, five points if they are on parole or probation, and one point for every police contact. In other words, every time a "chronic offender" is stopped by the police, a point is added to their score. "There are a lot of chicken-shit violations you can stop someone for," one sergeant explained during a ride-along. Every officer I asked said that stops still have to be constitutional, but that you could almost always find *something* to stop someone for. Another officer offered the following examples:

> Yesterday this individual might have got stopped because he jaywalked. Today he mighta got stopped because he didn't use his turn signal or whatever the case might be. So that's two points . . . say this individuals' walking, "Hey, can I talk

> to you for a moment?" "Yeah what's up?" You know, and then you just start filling out your card as he answers questions or whatever. And what it was telling us is who is out on the street, you know, who's out there not necessarily committing a crime but who's active on the streets. You put the activity of . . . being in a street with their violent background and one and one might create the next crime that's gonna occur.

The CID also created workups, referred to as "Chronic Offender Bulletins" on the individuals with the highest point values, which are distributed to officers at the beginning of their shifts. The goal of these bulletins is to give officers "situational awareness":

> There's two ways of looking at it. Either you start conducting your investigation to see if maybe there was a crime that had just been committed. Or, "We know who you are, you know, I just called you Johnny, I've never really met you before, but I know who you are now," so maybe it's put in [the guy's] mind, "Oh, they're on to me, they know who I am."

I readily saw this practice become a feedback loop in which officers were told to stop those with the highest point values, they stopped them, and that stop added another point to their score, further increasing the likelihood that they will be stopped again in the future, leading risk scores to become decoupled from crime or actual behavior.

Dragnet and directed surveillance are not mutually exclusive. Individuals with no direct police contact can be included in police databases merely through having a network tie to a person with direct police contact. For example, using Palantir's platform, LAPD officers can create a network graph where individuals with no direct police contact are drawn into law enforcement's "secondary surveillance network" as a function of their being colleagues, siblings, or lovers of the central person of interest.[4]

In light of these ethnographic findings, in the next section, I offer four provocations about the limits of law in governing police activity in the digital age.

III. PROVOCATIONS

A. Big Data Unsettles Underlying Legal Categories

Big data policing is shifting the ground under foundational legal concepts that govern police activity. The use of algorithms to predict criminal risk, for example, lays bare internal inconsistencies in basic legal principles such as individualized suspicion.

In principle, different police activities must meet different standards of suspicion. Reasonable suspicion is predicated on "specific" and "articulable"

[4] For an example of a secondary surveillance network in Palantir, see SARAH BRAYNE, PREDICT AND SURVEIL: DATA, DISCRETION, AND THE FUTURE OF POLICING 53–54 (2021).

facts "taken together with rational inferences from those facts."[5] However, police do not have to rely exclusively on their personal observations in their reasonable suspicion calculus. The Supreme Court held, for instance, that police observation in a "high crime area" can be a deciding factor with regard to reasonable suspicion or probable cause.[6] The Court never precisely defined what constituted a high crime area, and so it is reasonable to assume that the area inside a predictive policing box might qualify.

When big data—such as a predictive policing forecast—is combined with small data—such as traditional individualized suspicion based on particularized facts about a person observed at a given time and place—it effectively makes it easier for law enforcement to meet the reasonable suspicion standard. Some legal scholars have analyzed whether it is legally defensible for algorithmic predictions to play a role in law enforcement's suspicion calculus.[7] Andrew Ferguson writes, "If walking through a predicted red box changes my constitutional rights to be free from unreasonable searches and seizures, then a higher level of scrutiny might need to be brought to bear on the use of technology."[8] Although predictive models may be a preliminary factor in establishing reasonable suspicion, Ferguson contends that predictive policing forecasts are not sufficient to justify reasonable suspicion or probable cause because the data used in predictive policing forecasts is culled from previously observed circumstances with no direct relationship to the specific situation at the moment of the reasonable suspicion calculus.[9] In other words, individualized suspicion is incongruous with probabilities, because probability is not individualized.

However, the individualized/probabilistic distinction is a false binary. In most cases of individualized suspicion, the police officer does not actually observe someone committing a crime. Rather, they usually observe probabilistic indicia that they infer are associated with criminal activity. In that sense, what we might label "individualized suspicion" is actually a probabilistic assessment that criminal activity is afoot.

Consider this example: A person has a visible bulge, which is a common factor in establishing reasonable suspicion, because a bulge could be a gun. The police stop and search this individual. Is the stop defensible? Typically, an officer would not *know* that the person has a gun; rather, they see a bulge and are acting under a condition of uncertainty—is it a gun, is it a snack, is it an ostomy bag? A bulge becomes simply another way of articulating a probabilistic

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

[6] Elizabeth E. Joh, *Policing by Numbers: Big Data and the Fourth Amendment*, 89 WASH. L. REV. 35, 56 (2014).

[7] *E.g.*, *id.*; ANDREW GUTHRIE FERGUSON, RISE OF BIG DATA POLICING 77 (2017).

[8] FERGUSON, *supra* note 7, at 77.

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

assessment that Person X has a higher probability of being involved in criminal activity than Person Y. It starts with a statistical claim—at the population level, people with bulges are more likely than people without bulges to have guns, and guns are associated with criminal activity. Then comes the normative claim—because people with bulges are more likely than people without bulges to be involved in criminal activity, observing a bulge on a person contributes to a certain risk threshold, and it becomes legally defensible to stop them.

Legal doctrines can be thought of as routinized ways of accepting uncertainty and legitimating state power under those conditions. Acceptable police actions under uncertain conditions are then institutionalized in common law doctrine. From an organizational theory perspective, individualized suspicion—and the Fourth Amendment more broadly—is an institutionalized covering practice that does not reduce uncertainty, but rather renders it more acceptable.[10] Predictive policing is upsetting these settled ways of dealing with uncertainty and forcing us to articulate the normative logic of past practice.

In other words, at first it might seem like the challenge with predictive policing and reasonable suspicion is that predictive policing involves probabilities, whereas reasonable suspicion is individualized and particularized. But, actually, individualized suspicion is often a probabilistic assessment, too. Predictive policing does not simply erode the individualized suspicion versus probabilistic assessment binary; it sheds light on the fact that it was a false binary to begin with. Algorithmic suspicion forces us to rethink legal categories. Should the courts require a higher level of individualization because the source of the probabilistic determination is an algorithm, as opposed to a police officer's judgment? Humans were running such probabilistic assessments in their minds long before police used computers. There is nothing different about the logic involved in an officer seeing a bulge and conducting a quick mental probabilistic determination that criminal activity is likely afoot and a predictive algorithm saying Person X is more likely than Person Y to be involved in criminal activity. Except, crucially, you can argue about a suspicious bulge in court. A nonexpert can argue about the reasonable inferences a reasonable person can draw from seeing a bulge. But if the *algorithm* is uninterpretable, it is much more difficult to say whether it provides a fair or unfair basis for a stop.

B. Big Data is Different in Kind, Not Just Degree

Second, police are not simply scaling up data collection in the digital age; rather, different kinds of data are being produced. In particular, big data policing is programmatic, often suspicionless, and primarily operates below thresholds of reasonable suspicion and probable cause. Despite the fact that there is a

[10] John W. Meyer & Brian Rowan, *Institutionalized Organizations: Formal Structure as Myth and Ceremony*, 83 AM. J. SOCIO. 340, 340–41 (1997).

difference in kind—rather than just degree—old legal doctrine is still being laid on top of these data.

Take the automatic license plate reader, or ALPR, as an example. Some ALPR data is collected by the police and some by third parties, such as vehicle repossession agencies. Police access to aggregated ALPR data means that police have access to locational and temporal information on individuals who have never had any police contact. In that sense, the ongoing nature of license plate readings represents a proliferation of pre-warrant surveillance. Information is routinely accumulated, and files are lying in wait. Individuals' movements are codified as data, and once in a database, data points—such as the dates, times, and locations of an individual's vehicle picked up by an ALPR—can be marshaled as evidence *retroactively*, once that individual comes under suspicion.

Further, police can query datasets either deductively or inductively. Thus, if they are interested in knowing which individuals' cars were located near the scene of a crime, they can look up all license plates captured within a specified radius within particular time bounds. Conversely, if the police have a specific person of interest, they can search Palantir for all of the places and times an ALPR picked up that person's car. Once they are in a database, individuals can be surveilled repeatedly, which raises the question: Should new surveillance technologies facilitating the constant analysis and reanalysis of data be treated as "searches" subject to the Fourth Amendment? More fundamentally, is the Fourth Amendment even an appropriate tool for governing police activity in the digital age?

There have been thousands of judicial decisions interpreting the Fourth Amendment, and there are almost as many opinions among legal scholars about the Fourth Amendment's flexibility (or lack thereof) in the digital age. Some argue that an entirely different branch of law—administrative law—may be better-suited to govern police activity in the age of big data. For example, Daphna Renan argues that the traditional paradigm of Fourth Amendment law is *transactional*, focusing on the one-off interaction between law enforcement and a suspect. However, police surveillance in the age of big data is *programmatic*: it involves ongoing, cumulative, and sometimes suspicionless data collection and use.[11] Therefore, Renan suggests, administrative law may be a more appropriate legal framework than criminal procedure for governing cumulative surveillance.[12]

[11] Daphna Renan, *The Fourth Amendment as Administrative Governance*, 68 STAN. L. REV. 1039, 1042 (2016).

[12] *Id.* at 1058; *see also* Christopher Slobogin, *Policing and the Cloud*, NATIONAL CONSTITUTION CENTER (2017), https://constitutioncenter.org/digital-privacy/policing-and-the-cloud [https://perma.cc/9BBB-KCFS].

A related but distinct issue is that the police are increasingly using data originally collected by other, non-law enforcement actors. Function creep—the repurposing of data—is not well addressed by the Fourth Amendment. Although it is not an entirely new concern,[13] what *is* new about the digital age is that individuals leave so many more seemingly innocuous digital traces in the course of their everyday lives that can now be linked together using big data systems, and thus grant law enforcement a level of insight that almost certainly would have been considered an intrusive investigation in years past.

C. Big Data has a Life Course

Third, relying on extant legal mechanisms like the exclusionary rule means using what is meant to be a check on state power at one point in time and space, whereas data is fundamentally social and, as such, has a life course.

It remains an open question whether predictive policing will lead to revolutionary change in standards of reasonable suspicion. Yet the courts will likely accept the outputs of predictive models as one of the many defensible factors that can be taken into consideration for suspicion in the age of big data—*that is, if the cases even make it to court.* As the sociolegal scholar Issa Kohler-Hausmann points out, "trials have gone the way of the dodo bird."[14] Many people accept plea deals, for instance, for a raft of reasons: they can't afford to take time off work for a trial, they can't risk getting a longer sentence than the one offered by the plea (even if they insist on their innocence), and so on. But if they don't go to trial, the basis for their arrest will not be examined in a court of law. It is functionally impossible to consider how often courts might rule against the legality of algorithm-informed police contact, because they so rarely have the chance to do so. As Kohler-Hausmann explains, "Fourth Amendment jurisprudence is built on the premise that substantive rights . . . are secured by the mechanism of excluding unlawfully seized evidence and arrests . . . [but] [t]he overwhelming amount of police work is low-level enforcement activity, not serious violent felony arrests."[15] In other words, we can't simply rely on the exclusionary rule, which says that the police cannot use particular pieces of information in trial, as *the* opportunity for defendants to invoke their procedural rights; challenge the legality of a police stop, search, or seizure; or establish and exclude evidence as inadmissible, because law enforcement use of big data is so infrequently scrutinized in an adversarial trial context.

That is why it is important to structure the police's use of data and power in a way that protects the underlying interests that we thought were at issue in the

[13] Harold J. Krent, *Of Diaries and Data Banks: Use Restrictions Under the Fourth Amendment*, 74 TEX. L. REV. 49, 50 (1995).

[14] ISSA KOHLER-HAUSMANN, MISDEMEANORLAND: CRIMINAL COURTS AND SOCIAL CONTROL IN AN AGE OF BROKEN WINDOWS POLICING 258 (2018).

[15] *Id.*

Fourth Amendment to begin with: protection from arbitrary, unjustified state intrusions into people's lives.[16] In Kohler-Hausmann's words, "If we are serious about bringing . . . police activity into line with substantive legal principles governing stops, searches, and seizures, then we must innovate other political and organizational mechanisms to do so."[17] We should look not only to law, but also to political incentives and more mundane organizational practices to protect individual rights.

In more sociological terms, cut-offs do not make sense because they do not reflect the way the world works. Data is social and relational, and it has what we can call a "life course."[18] So although the law is better equipped to look at specific cut-off points, we have to look at the whole sequence of events in determining reasonable suspicion. For example, the exclusionary rule says you cannot admit illegally obtained evidence in court. But law is only considering the data in its mature life-course stage. We need to think about the nascent data, produced in social contexts and connected to other data points across time and space. If the underlying data are flawed, one might argue, the visible data—the predictive box or the Operation LASER score—may lose its defensibility.

D. Big Data Creates New Opportunities for Parallel Construction

Fourth, unfettered big data policing creates new opportunities for information asymmetries and can threaten due process through a long-standing practice called "parallel construction." Parallel construction is defined in a training slide from the Drug Enforcement Administration ("DEA") as the use of normal investigative techniques to recreate the information provided by the Special Operations Division of the DEA ("SOD").[19] It is the process of building a separate evidentiary base to conceal how an investigation began because it involved informants, warrantless surveillance, or other inadmissible evidence. Or, in the words of Nancy Gertner, a former U.S. federal judge, parallel construction is "a fancy word for phonying up the course of the investigation."[20]

For example, say law enforcement learned about a crime through an intelligence surveillance program or platform it wishes to keep secret. Law enforcement could claim the information came from a confidential informant or

[16] *Id.* at 259–60.

[17] *Id.* at 259.

[18] Karen E. C. Levy, *Relational Big Data*, 66 STAN. L. REV. ONLINE 73, 75 (2013), https://www.stanfordlawreview.org/online/privacy-and-big-data-relational-big-data/ [https://perma.cc/XX6N-C4SY].

[19] *Dark Side: Secret Origins of Evidence in US Criminal Cases*, HUMAN RIGHTS WATCH (Jan. 9, 2018), https://www.hrw.org/report/2018/01/09/dark-side/secret-origins-evidence-us-criminal-cases# [https://perma.cc/EHP3-WMBS].

[20] Joe Kloc, *DEA Investigated for Using NSA Data for Drug Busts*, DAILY DOT (May 2, 2020, 4:06 PM), https://www.dailydot.com/layer8/dea-sod-nsa-snowden-investigation/ [https://perma.cc/Q6Y8-ZHDF].

follow the individual of interest until they neglect to use their turn signal, at which point they can conduct a pretextual stop and recreate the investigative trail such that it appears the traffic stop was the point at which the investigation began. If they end up arresting the individual, nothing admitted as evidence would say anything about the surveillance strategy. "It's just like laundering money—you work backwards to make it clean," said one former DEA agent.[21]

The use of big data for parallel construction can increase information asymmetries—and therefore power asymmetries—in the criminal legal system. In essence, parallel construction helps law enforcement to circumvent Fourth Amendment protections, ultimately violating defendants' constitutional right to a fair trial by obscuring the means by which they came under law enforcement surveillance. By misleading the court about what happened, the court, too, is disempowered by parallel construction. It cannot determine what really happened and whether or not it was legal.

Moreover, if surveillance strategies are not disclosed, we may underestimate the extent of privacy and civil liberties violations occurring, judges will not have the opportunity to evaluate the constitutionality of law enforcement activities, and defendants may be denied the opportunity to know about and access exculpatory evidence in violation of pretrial discovery rules.[22]

Even though the concept itself is not new, as the number of ways police use surveillance technologies and new sources of data increases, so too do the possibilities for parallel construction. For example, at a surveillance industry conference I attended, a member of Palantir's legal counsel explained how the platform could be used to knit together circumstantial evidence into a comprehensive picture. Whereas it is relatively rare to find a "smoking gun," he explained, Palantir users might be able to build up a sequence of events that could have been impossible for prosecutors to construct in the past. By integrating data into a "single ontology," you can draw connections between actors and depict a coherent scheme. On the one hand, this could be incredibly useful to investigators who previously operated in siloed jurisdictions with limited access to historical information on potential suspects. But on the other hand, officer hunches that would be insufficient grounds for obtaining a warrant can now be retroactively backed up using existing data, and queries can be justified *after* data confirm officer suspicions.

Palantir has tried to address some concerns about abusive uses of their platform. The main Palantir protection against abuse that came up in my interviews was the "immutable audit log" that shows who accessed what data

[21] John Shiffman & Kristina Cooke, *Exclusive: U.S. Directs Agents to Cover Up Program Used to Investigate Americans*, REUTERS (Aug 5. 2013, 4:19 AM), https://www.reuters.com/article/us-dea-sod/exclusive-u-s-directs-agents-to-cover-up-program-used-to-investigate-americans-idUSBRE97409R20130805 [https://perma.cc/A47U-2KML].

[22] HUMAN RIGHTS WATCH, *supra* note 19; Brady v. Maryland, 373 U.S. 83, 87 (1963).

when. However, just because Palantir's system has a built-in audit capacity does not mean that their clients use it in practice. I was unable to find a single instance, for example, in which the LAPD conducted an audit of their logs in Palantir.

To be sure, I did not observe parallel construction occurring during my fieldwork. In this project, I did not follow cases into criminal legal processing after police contact. The point here is simply to suggest that big data increases the *opportunities* for parallel construction to take place. Future research can investigate the extent to which this does or does not occur.

CONCLUSION

Before big data analytics became a new tool in the law enforcement arsenal, there were flaws in the criminal legal system. So too are there cracks in the laws governing police activity. Counter to the Silicon Valley rhetoric that big data is a disruptive force transforming everything as we know it, it is useful to consider big data as the injection of a new technology, a dye that illuminates the cracks in the system and brings fault lines into stark relief. Algorithms serve to formalize and codify the social practice of policing in ways that call into question the fundamental legal frameworks underpinning longstanding checks on police power.