2018

The Legal Risks of Big Data Policing

Andrew G. Ferguson
American University Washington College of Law, ferguson@wcl.american.edu

33-SUM Crim. Just. 4

Criminal Justice
Summer, 2018

Feature
Andrew Guthrie Ferguson [a1]

THE LEGAL RISKS OF BIG DATA POLICING

The future of law enforcement is being shaped by new technologies. Today, on the streets of major cities, algorithms forecast areas of predicted crime, risk models create lists of possible suspects, and social network analysis targets criminal groups for increased surveillance. In practical effect, technology is changing where police patrol, who they target, and how they do their jobs.

Yet, despite this rapid advancement, the law has remained decidedly stuck in the past. In fact, the legal risks involved in these innovations remain largely unexamined by the legal profession. In the near future, lawyers will need to take on new roles in responding to these policing strategies. Prosecutors and defenders will need to litigate more aggressive digital surveillance techniques. Judges will be required to retrofit ancient constitutional doctrines to meet new technological challenges. And lawyers for the entrepreneurial engines of growth--the companies--will need to conduct risk assessments about the litigation dangers arising from these new surveillance capabilities.

***5** This article seeks to examine the changing law enforcement reality with an eye toward legal risk. The rise of big data policing creates real opportunities and substantial dangers, and so far the legal profession (as an organizing force) has not played a central role, generally deferring to technology innovators, police administrators, and civil rights groups to drive the debate. This should change. After all, lawyers and their families live in these policed communities, will be litigating the issues in criminal court, and will need to provide important advice to companies thinking through some of the litigation risks involved in developing and implementing new technologies. As citizens, advocates, and counselors, the rise of big data policing provides a chance for lawyers to engineer the future balance between security and liberty.

THE GROWTH OF BIG DATA POLICING

In more than 60 American cities, police are using some form of predictive policing to deter crime. (David Robinson & Logan Koepke, Upturn, *Stuck in a Pattern: Early Evidence on "Predictive Policing" and Civil Rights* 3-5 (2016).) Police departments have partnered with small start-up companies and academic enterprises to forecast the places most likely to be the location of a crime. The general theory behind predictive policing is that particular types of crime can be identified by studying past crime patterns. Some predictive policing algorithms only rely on past criminal incidents, day, time, and place, while others add in more complex variables like the time of year, weather, and particular local factors (fairs, football games) and yet other models study fixed structures that might encourage criminal activity (bus stops, liquor stores) providing the cover for loitering and/or the targeting of victims. (*See generally* Andrew Guthrie Ferguson, *Policing Predictive Policing*, 94 WASH. U. L. REV. 1113 (2017).) In a predictive policing jurisdiction, the crime numbers are crunched and spit out into usable maps that can identify particular areas of possible crime so that police can patrol those areas. The goal is "to predict and deter" under the logic that if the risk forecast is accurate, the police presence will deter the potential criminal actor from following through on his criminal plan.

Some cities like Chicago and Manhattan also have begun using predictive analytics to identify people more at risk of being involved in violent crime. By analyzing past criminal arrests, convictions, age, and other things like gang activity or being the victim of violence, these cities have set up programs to identify and intervene in lives of these "at-risk" individuals. (Jeremey Gorner, *Chicago Police Use "Heat List" as Strategy to Prevent Violence*, CHI. TRIB. (Aug. 21, 2013).) So, for example, in Chicago, the "Strategic Subjects List" creates a rank-ordered list of the people in Chicago who are most at risk at being either the perpetrator or victim of a violent crime. Each identified person is given a threat score from 1 to 500+, with the police attention and focus being on those with the highest scores. (Josh Kaplan, *Predictive Policing and the Long Road to Transparency*, S. SIDE WKLY. (July 12, 2017).) The theory behind what has colloquially become known as the "heat list" is that risk is not spread equally in a society. In fact, the risk of violence clusters among certain groups and, thus, police resources should be directed at those risky groups. The theory underlying the heat list arises from a recognized pattern of reciprocal violence. When one person is shot, then that person's friends (or fellow gang members) might seek revenge by shooting the perpetrators, which, in turn, will create more acts of violence. The predictive element comes from the pattern of revenge, and the solution of predictive policing systems is to intervene to break that cycle of violence. These "focused deterrence" tactics usually involve police officials (along with some social services representatives) visiting the targeted individual to detail his risk score and the need for him to remove himself from this cycle of violence. Sometimes the individuals are called into community meetings, and sometimes the police literally knock on their door to give the warning, but the message is the same; You are being watched and you are at risk. By utilizing predictive analytics, police believe they can get a better sense of the patterns of violence in a community.

Groups of suspected criminals also are being watched. The Los Angeles Police Department (LAPD) has partnered with the private data company Palantir to build a social network investigation system to monitor gangs and chronic offenders: (Mark Harris, *How Peter Thiel's Secretive Data Company Pushed into Policing*, Wired (Aug. 9, 2017).) In big cities facing cross-jurisdictional crime problems, the ability to monitor large areas and large groups has meant a new focus on data collection and analysis. Police are tasked to contact and monitor "chronic offenders" (identified by having a high-risk score resulting from past involvement in the criminal justice system). (*See generally* Sarah Brayne, *Big Data Surveillance: The Case of Policing*, 82 AM. SOC. REV. 977 (2017).) Police contact these chronic offenders and fill out field interview cards--data about place, who the targets are hanging out with, their car, home, etc. Then all of this information is inputted into the growing police database that can be used to link criminal associates (and others), to visualize patterns of criminal activity, and to investigate crime. (*See id.*) A shared address can link different groups just as easily as a shared car or cell phone number. For investigative purposes, this growing database can provide clues never before recorded or utilized. As a tool, this strategy both offers a measure of social control on those most at risk, but also provides a valuable investigative resource if there is a crime that needs to be solved.

Police are not only monitoring data, but also watching us. In big cities like Manhattan, linked video cameras feed into a central command center that both has real-time observation capabilities but also can rewind the tape if an incident should occur. Police body cameras are building a library of lived experience of citizens on the streets. Automated license plate readers record the location of our cars. Shotspotter audio collection systems listen for gunshots. Chemical detection devices sniff our scents. Biometric collection is growing with more than half of Americans' facial images now in the systems that can be searched by facial recognition software. DNA, iris scans, and other pattern-matching systems are being built to identify people from public surveillance. The ordinary senses of police power are being supercharged by advanced ***6** technologies. And all of these technologies are capturing personal data with an eye toward future law enforcement use.

And, of course, we creatures of consumer convenience are creating wonderfully revealing data trails exposing every step we take, every purchase we make, every question we have for Google's search algorithm. As we go about our daily lives, we are tracked by the smartphone in our pocket. In the future, tech-connected criminals will leave their smart homes, hop into their smart cars, with their smart Internet of Things--enabled devices tagging along, and put shoe-leather detectives out of business because it will be just too easy to reveal where they went and what they did there. How police will use smart devices to snitch on our most private activities is only now being litigated.

Currently, these big data policing technologies are in their early stages. Cities have begun experimenting with different innovations, but it still remains a fragmented reality. This fragmentation is exacerbated by the fact that law enforcement itself is a fragmented profession with upwards of 17,000 different law enforcement agencies in America. (Barry Friedman & Maria Ponomarenko, *Democratic Policing*, 90 N.Y.U. L. REV. 1827, 1843 (2015).) That said, the technology is rapidly improving and getting cheaper and easier to adopt. As one can see, the technologies portend a radical rebalancing of privacy interests in the name of public safety. More and more cities have expressed interest in using predictive policing, more police are tracking the Internet of Things, and the rules of how these technologies will be used in court, as of yet, are undecided.

THE LEGAL RISKS

Big data policing creates a new set of legal risks. The possible issues associated with privacy invasions, constitutional rights, municipal liability, corporate health, and how any of these technologies get litigated in criminal court are myriad and largely unmanaged. In fact, in many small start-up companies, venture capital firms, and academic innovation labs, lawyers are not even in the room. And lawyers should be in the room, if only because the questions are endlessly fascinating.

Take, for example, a few of the open constitutional questions. Do current forms of mass surveillance fall outside of the Fourth Amendment? (Stephen Rushin, *The Judicial Response to Mass Surveillance*, 2011 U. ILL. J.L. TECH. & POL'Y 281, 285-86 (2011).) Does the current understanding that citizens have no reasonable expectation of privacy traveling from point A to point B still hold when city-wide surveillance also can track individuals using facial recognition going from point A to point Z (including your trips to the health clinic, your client's home, and the local political resistance group)? Or how do the growing networks of "smart" effects that make up the Internet of Things fit a Fourth Amendment framework? (Andrew Guthrie Ferguson, *The Internet of Things and the Fourth Amendment of Effects*, 104 CAL. L. REV. 805, 823 (2016).) Is the data coming from smart health devices, cars, or other objects protected by the words of the Framers? We do not know yet. Nor do we know how courts will react to the privacy-invading, but law enforcement--enabling use of aggregated data sources. Cameras that catch everything offer game-changing investigative promise, but also rework existing conceptions of privacy. Add in facial recognition capabilities or social media search capabilities, and you have a truly powerful surveillance power without any clear Fourth Amendment guidance.

Or what about the question of how courts should evaluate the legal weight of a predictive policing tip? If an algorithm tells a police officer to go to a particular block at a particular time to be looking for a particular crime and the officer sees something suggestive of that crime, wouldn't the prediction naturally impact reasonable suspicion? (*See* Andrew Guthrie Ferguson, *Predictive Policing and Reasonable Suspicion*, 62 EMORY L.J. 259, 304 (2012).) But what if we have no way to judge the accuracy of the algorithm, and thus a faulty forecast could be changing constitutional protections in certain parts of the city? Or what if a police officer in Chicago stops a "suspect" who has the highest possible score from the "heat list"? Won't that threat score change how the officer approaches the person, or whether she uses force or interacts with him? And, again, what if we have no idea about the accuracy of the underlying algorithm that causes the elevated threat score? (*See generally* Andrew Guthrie Ferguson, *Big Data and Predictive Reasonable Suspicion*, 163 U. PA. L. REV. 327 (2015).)

Or what about the legal issues from the city's perspective? Cities need to balance calls for public safety with community trust. Do new big data technologies add or detract from that trust-building project? Do city-wide camera systems send a reassuring or threatening message? And does it depend on the community being targeted? Privacy concerns do not impact just individual liberty but can result in significant financial penalties in the form of lawsuits and challenges to new surveillance practices. If a police data system gets hacked, if a facial recognition system creates a false match, or if a citizen FOIAs for all of the automated license plate readings to track his estranged wife, the municipal equities get complicated quite quickly. How should city attorneys mitigate the legal risks before a lawsuit is filed?

The risks are even greater for small start-up companies that cannot afford the litigation costs of lengthy lawsuits or even sometimes the transparency required for court cases. If you run a predictive policing company whose product is essentially a proprietary algorithm, what happens when the judge demands you reveal the process in open court? If you are a police body camera company, how do you protect confidential matters like juvenile defendant footage or private health information? If you are a data company (and almost all tech companies are now data companies), how can you ensure consumer privacy from growing law enforcement demands? These legal issues should be asked by companies at the front end but rarely are because of the lack of lawyers involved in the design or engineering stage.

Finally, if you are a defense lawyer, prosecutor, or judge, how do any of these technologies play out in court? What are the evidentiary limitations on algorithms? What are the reliability requirements for admissibility of social network patterns? What are the Brady protections built into big data investigative systems? What if the algorithms demonstrate *7 a racial bias? The list goes on far longer than any answers have been developed.

RESPONDING TO LEGAL RISKS

Lawyers manage risk, and the growing risks of big data policing create new opportunities for curious lawyers interested in developing along with the technology. Part of the reason for writing my book--*The Rise of Big Data Policing: Surveillance, Race, and the Future of Law Enforcement*--was to encourage engagement about these questions at the front end. There are not enough lawyers thinking through the litigation risks or even generating the litigation, despite the growing influence of the technologies.

The need for interested lawyers remains at every stage of the process. For example, lawyers for start-ups need to think through the legal and ethical challenges of predictive analytics, artificial intelligence, and new surveillance systems. Lawyers can assist in building a legal and moral accountability by identifying future legal and ethical risks. If these legal questions are not addressed at the design or engineering stage, the difficulty in fixing them after implementation only increases (and grows more costly).

Lawyers for cities need to think about the data being collected from their citizens. What responsibility do city lawyers have to protect the public's private data or privacy? Who owns the data collected on city streets and what can be done with the data and by whom? Who profits from the collection? Who is informed about breaches? What happens to the data when the technology becomes obsolete or the private company goes bankrupt? These issues about data collection, use, and storage must be written into city contracts and thought through with an eye for future litigation.

Lawyers in court need to think about the fairness concerns beyond just evidentiary admissibility. How are new technologies reifying racial bias or economic inequality? How much trust should we put in invalidated systems? How can juries evaluate reliability or accuracy or fairness? How can lawyers cross-examine an algorithm? And how do judges have the capacity to sort through the legal questions without the help of educated lawyers or experts.

These are the challenges of a big data future. These tasks of setting rules, guidelines, and principles around how to balance the risks and rewards of new technology should fall to the legal profession. While the American Bar Association (ABA), the American Law Institute (ALI), and the National Institute of Justice (NIJ) among other national thought leaders have pushed the conversation forward, more can be done to engage the profession. Law schools and legal clinics can play a larger role. Law students can see the need for technological fluency. Law firms should see it in their financial interest to invest in answering some of these hard questions and to support the development of civil society responses to these risks. A national conversation should begin and be led by the legal profession.

When it comes to policing, privacy, and criminal justice, lawyers should play the central role in setting forth best practices surrounding surveillance technologies. There is an urgent need for an honest broker to be able to sort through the

competing demands of civil rights and public safety. There is an urgent need to decide who should have access to the data, who should draft the rules, and who can profit from them. The legal profession should fund and promote a national task force to address these problems. The legal profession, with the support of academia, police leaders, technologists, and civil libertarians, should be at the table to debate the challenges and the path forward.

Lawyers possess the training and skill to foresee the legal risk ahead. If educated about the challenges of how big data policing is changing policing, lawyers from all walks of life can join the conversation. After all, the goal of any predictive risk assessment, like the job of any good lawyer, is to foresee future risk, and so lawyers should embrace that predictive mindset and begin designing the future of big data policing today.

Footnotes

a1 **ANDREW GUTHRIE FERGUSON** is a professor of law at the UDC David A. Clarke School of Law and author of *The Rise of Big Data Policing: Surveillance, Race, and the Future of Law Enforcement* (2017). Twitter @ProfFerguson

33-SUM CRIMJUST 4

End of Document