What Big Data Means for the Fourth Amendment

Author : Elizabeth Joh

Date : January 6, 2016

- Andrew G. Ferguson, *Big Data and Predictive Reasonable Suspicion*, 163 **Univ. Penn. L. Rev**. 327 (2015).
- Michael Rich, *Machine Learning*, *Automated Suspicion Algorithms*, *and the Fourth Amendment*, **__ Univ. Penn. L. Rev.** __ (forthcoming 2016), available at SSRN.

Hear the term "big data," and the police are not likely to be the first word that comes to mind. Whether or not you are familiar with the term, the vast quantities of digitized information available today—and the data analytics that are applied to it—already shape your world. The movie recommended to you by Netflix, the date you chose on OkCupid, or the ad you saw on your Facebook feed are all the result of the pervasiveness of big data. That same big data revolution is coming to policing. The NYPD operates a "domain awareness system" that links license plate reader data, "smart" cameras, law enforcement databases, texts of 911 calls, and radiation sensors information from around the city. Police departments in Seattle and Los Angeles are piloting predictive policing software that directs officers to places where crime is most likely to happen in the future. Other law enforcement agencies are considering the adoption of social media software that sifts through tweets, likes, pins, and posts for potential on-line threats. To be sure, the police have always relied upon large quantities of data, but the promise of "big data" lies in its enormous volume, its reach, and the application of sophisticated computer analytics.

In response, there is a small but important emerging scholarship that addresses some of the difficult questions posed by the use of big data by the police. In two recent pieces, both Andrew G. Ferguson and Michael Rich address these issues especially well. While each focuses on different aspects of big data use, and each comes to different conclusions about the Fourth Amendment implications, this pair of articles introduces an evolving set of concerns that should be incorporated into every criminal procedure scholar's current knowledge.

The legality of an officer's decision to engage in a so-called stop-and-frisk depends on a finding of individualized reasonable suspicion. That's the law, of course, but defining the content of reasonable suspicion has always been tricky. As many law professors have lamented, the *Terry* standard is malleable to the point of being meaningless. To make matters worse, the Supreme Court has repeatedly emphasized that reasonable suspicion is neither a quantifiable nor a scientific concept.

Enter big data. What if the police use big data programs as the basis for stops or frisks? What if, for instance, an officer is alerted by a piece of software that trawls through millions of pieces of data to predict that a person is highly likely to commit a violent crime (based on an algorithm of his social connections, criminal history, and social media posts) in the very place the officer finds him (based on an algorithm using historical crime data)? Should a court uphold a stop and frisk that subsequently results in the discovery of an illegal firearm?

As both Ferguson and Rich point out, the Fourth Amendment fails to provide easy answers. This difficulty arises in part because traditional individualized suspicion is itself a fuzzy concept. And the most relevant analogies to big data programs are limited. For instance, we might compare big data programs to drug-sniffing dogs, since both are instruments for turning raw data into assessments about criminal suspicion. But direct application of the Supreme Court's drug dog cases poses problems, however, since the programming of a predictive algorithm is far more complex than a dog's sniff to both the officer who relies on its predictions and the judge who evaluates it (Rich, Pp. 60-62.)

Moreover, Ferguson and Rich arrive at different conclusions as to whether big data alone would provide reasonable suspicion for a stop. Ferguson, who uses a broader definition of big data than does Rich, contends that reliance upon big data could provide a far greater quantity of detailed information than any individual officer or informant coming from a "small data" perspective ever could. Rich, who focuses on predictive automated suspicion programs, argues instead that big data alone is insufficient for Fourth Amendment suspicion. A prediction made by software is incapable, contends Rich, of providing a true totality-of-the-circumstances assessment as required by Supreme Court precedent.

These differences matter less than the broader insights about big data that both Ferguson and Rich identify. The use of big data may provide distinct advantages over traditional policing. Contrary to longstanding concerns that individual officers use proxies like race, class, and neighborhood as the basis for suspicion, big data can bring more accuracy and precision to policing (Ferguson, P. 389.) And if suspicion is increasingly based upon data—rather than human intuition—then we might find a greater emphasis on accountability and transparency in policing as a result (Ferguson, P. 393.)

On the other side of the ledger, big data's promise of objective analysis may be misleading if, for example, its results rely on mistaken inputs. But finding mistakes in these enormous databases, often handled by countless persons and analyzed by "black box" algorithms may be near impossible. To make matters worse, as Rich points out, the Supreme Court's expansion of the good faith doctrine in cases like *Herring v. United States* (2009) establishes enormous obstacles for defendants challenging big data accuracy. How easily can a defendant demonstrate that a stop or frisk in his case was based on "deliberate, reckless, or grossly negligent" misconduct, or "recurring or systemic negligence"? Not very, argues Rich, and as a result, we may end up with a system in which "bad data and benign neglect could flourish" (P. 66.)

To be sure, those programs that will strain current Fourth Amendment doctrine most severely—predictive programs that will alert police to criminally suspicions persons—have not yet become part of ordinary policing. But both Ferguson and Rich agree that these programs are coming, and soon. And as with so many new policing technologies, the law is lagging behind (Ferguson, P. 410.) These two excellent pieces demonstrate the need to think about the inevitable widespread use of big data by the police in a systematic and reflective way, before the reality on the ground gets too far ahead of the law that is meant to govern it.

Cite as: Elizabeth Joh, *What Big Data Means for the Fourth Amendment*, JOTWELL (September 15, 2015) (reviewing Andrew G. Ferguson, *Big Data and Predictive Reasonable Suspicion*, 163**Univ. Penn. L. Rev**. 327 (2015) and Michael Rich, *Machine Learning*, *Automated Suspicion Algorithms*, *and the Fourth Amendment*, __ **Univ. Penn. L. Rev.** __ (forthcoming 2015), available at SSRN.), http://crim.jotwell.com/?p=1018.