

Freeing Up Data Science in Violent Crime Prevention

The Federal Bureau of Investigations, operating under the United States Department of Justice, is tasked with enforcing federal law, gathering evidence in the course of investigating and, ideally, preventing violent crimes before they occur. Focusing on the latter amidst our technological age, the agency—with its 50-plus field offices and over 30,000 employees—has a wealth of investigative resources, including numerous crime data sources, with which to assist local law enforcement and communities across the country. However, while rates of violent crime domestically have fallen overall in recent history, homicide rates in many cities have continued to climb, and the countrywide percentage of homicides going unsolved has increased, as illustrated by the Murder Accountability Project (MAP); Thomas Hargrove of MAP discovered over seven years ago that the FBI has a preponderance of unsolved homicides data going back to 1980 (Kolker, 2017). Herein lies an opportunity for the FBI to use data science to improve the way violent crimes are investigated and prevented.

In order for the FBI to solve the core business problem of rising homicide rates in many cities and a corresponding increase in unsolved, potentially-linked murders, we can consider the following questions:

- How can the agency capture more violent crime data and facilitate analysis, then disseminate those data, findings and methods to the proper law enforcement departments in order to guide local investigations?
- What data science methods can reasonably be employed/encouraged by the FBI, local law enforcement, and the private sector to predict and prevent homicides rather than principally reacting to them?

Answering these questions will require multiple levels of decision-making, with both persuasion and coordination needed to reap the benefits of data science.

The decisions primarily for FBI directors to make will surround (1) how and when to pursue investigations tipped by data analysis, and (2) how to compel local law enforcement to both contribute data upward as well as utilize the consolidated FBI crime data. This will require coordinating with chiefs of local police and mayors, whose priorities include decreasing violent crime, as well as getting input from the private sector: should these bubbled-up crime data be made more readily available publicly to

organizations such as MAP, with potential rewards for identifying the presence of serial killers? Currently, most FBI data require formal, lengthy Freedom of Information requests, with often limited data—and often scanned copies—available through its records management system (FBI website, 2017). Further, should data collection from local agencies continue to be largely voluntary, or should more pressure be applied to compel participation in FBI record-keeping? And lastly, should the FBI (re)allocate resources to building data science teams as well as training local law enforcement in basic analytics? The FBI has over 50 separate field offices in or near key cities, but like any government agency, acquiring funding for new initiatives requires political pleading, while local law enforcement tends to rely on “paper-based, scattered, and siloed” techniques (Kolker, 2017). Unfortunately, these decisions are likely to be met by typical bureaucratic biases, such as allegiances to the status quo and special interests; “tunnel vision” and “linkage blindness” during investigations, as Hargrove laments (Kolker, 2017); and anchoring to conclusions (e.g., as Boston police has to serial killer theories [Anderson, 2017]).

Yet there are both simple and complex analytical methods that quickly assist with investigation decision-making. While the FBI has published exploratory data analyses on homicides and victim demographics (Morton et al., 2017), the consolidated homicide data generally do not make their way back to local investigators in their geographically-expanded form. As Hargrove points out, even simple queries of open crime files beyond city or county borders, filtered by victim profile, are still novel methods that can quickly illustrate trends (Kolker, 2017). For example, seeing a preponderance of one type of homicide across an expanded region can lead investigators to cross-reference homicides of known criminals with similar streaks elsewhere.

Meanwhile, as the FBI seeks to grow its technological capabilities, analysts internally as well as in the private sector can identify links between unsolved crimes by developing more complex machine learning algorithms and statistical analyses. Hargrove, at the time a journalist, developed one such cluster analysis in 2010 that indicated there likely was a serial killer in Gary, Indiana; however, his tip to police went unrecognized before they captured the murderer several victims later in 2014—the link still unknown at the time (Kolker, 2017).

And so begins a feedback loop between the three levels as the data collected and disseminated by the FBI increase in scope, with confidence and participation in the system continually growing. The loop, however, begins with FBI directors recognizing the predictive potential of algorithms and then getting commitment from and guiding local agencies, particularly in cities with high crime rates. It also requires convincing FBI leadership and the executive branch’s budgetary decision-makers to allocate resources to data science, as well as consolidating and expanding access to master crime datasets; in the

world of government, these likely require political lobbying or legislative action, in addition to the cultural and process changes between federal, local and private entities.

With the embrace of data science can come a reversal to the trends of violent crime rates in many cities, as well as to the overall rate of violent crimes going unsolved. The nature of law enforcement could become less reactive and more proactive and preemptive, by picking up on patterns in data to prevent future violent crimes. In the process of expanding data-driven decisions, at a minimum we expect an improvement to dialogue and cooperation between federal and local law enforcement, as well as with private analytics firms; as Thomas Hargrove has shown, opening up the FBI's crime data allows for testing of hundreds of computer algorithms with ever-improving predictive power (Kolker, 2017). While this effort may increase the risk of pursuing false flags and privacy concerns for families who experience tragedy, it nonetheless improves agency coordination and investigative capabilities and ultimately saves lives by *preventing* more violent crime.

Works Cited

Anderson, Travis, "Evans refutes serial killer theory for bodies found in waterways," *Boston Globe*, 24 May 2017, website, accessed 3 July 2017, <https://www.bostonglobe.com/metro/2017/05/23/bpd-commissioner-refutes-serial-killer-theory-for-bodies-found-waterways/51A35WMn4zHhOjcvJ3htXO/story.html>

"Freedom of Information/Privacy Act," FBI website, accessed 3 July 2017, <https://www.fbi.gov/services/records-management/foipa>

Kolker, Robert, "Serial Killers Should Fear this Algorithm," *Bloomberg Businessweek*, 8 February 2017, <https://www.bloomberg.com/news/features/2017-02-08/serial-killers-should-fear-this-algorithm>

Morton, Robert J., Jennifer M. Tillman, Stephanie J. Gaines, "Serial Murder: Pathways for Investigations," 2017, National Center for the Analysis of Violent Crime, FBI website, accessed 3 July 2017, <https://www.fbi.gov/file-repository/serialmurder-pathwaysforinvestigations.pdf/>

Murder Accountability Project, website, accessed 3 July 2017, <http://murderdata.org/>

"Records Management," FBI website, accessed 3 July 2017, <https://www.fbi.gov/services/records-management>

"What We Investigate," FBI website, accessed 3 July 2017, <https://www.fbi.gov/investigate>