

This document was produced based on notes taken during the Predictive Policing workshop of the Data & Civil Rights conference. This document represents a general summary of the discussion that took place. Not all attendees were involved in every part of the conversation, nor does this document necessarily reflect the views and beliefs of individual attendees. All workshop participants received workshop materials prior to the event to spark discussion. The primer can be found at: http://www.datacivilrights.org/pubs/2015-1027/Predictive_Policing.pdf

Workshop Discussion Notes: Predictive Policing

Overview

This workshop focused on the increasing use of data analysis tools for Predictive Policing by law enforcement agencies. Many police departments now have formal or informal programs that use various data tools to inform policing activities, from the identification of “hotspots” to individuals of note. Participants in the Predictive Policing workshop came from a wide variety of backgrounds; this included law enforcement officials, community advocates, technologists, and academic researchers. During the workshop, participants were given the chance to explain their investment in the area of predictive policing and to frame what they saw as some of the central concerns with the use of data analysis in crime prediction, deterrence, and response. Among workshop participants, questions about algorithmic transparency and efficacy were raised, the issues of police discrimination were flagged, and the pervasive cultural specter of “Minority Report” was lamented. The session primarily organized the discussion around the inputs, actions, and evaluation of predictive policing methods. The workshop discussed the current state of Predictive policing, the imagined future of the practice, and a set of possible interventions.

Themes and discussion topics

The Current Uses of Predictive Policing

The conversation began by describing how predictive policing methods are currently being integrated into the larger system of law enforcement. Most of what takes place currently falls under the rubric of data mining or predictive analysis, but has been rebranded “predictive policing” to avoid negative public opinion. One type of this data mining is focused on locations: identify a place, and send more police to that place or focus a single officer’s attention on a small area or block. This can be thought of as the newest incarnation of “hot spot” policing, which has been in use by police for quite some time. The motivation for and justification of these practices is deeply entwined with economic concerns. Many of the techniques and technologies themselves have been adapted from commercial settings (where actors like Walmart or

Amazon use analytics to operate a just-in-time business model). Some argue that in commercial settings, experimentation with new, expensive methods can be considered a worthwhile investment, but when it comes to the allocation of limited public dollars, these methods and their contribution to public safety may not be worthwhile, particularly given their limited proven connection to increased public safety. Support for predictive policing methods also intersects with budgetary arguments at a political level between police departments and the municipalities or cities where they receive funding from: the norm in policing for a long time was that police can respond to, but not predict crime. The promise of predictive policing technology is that it enables police to use technology to prevent crime. This, in turn, leads to safer neighborhoods. This argument is leveraged in political discourse to argue for increased budgets for local departments so that they can purchase tools; conversely, depriving them of that funding threatens to reduce neighborhood safety. Simultaneously, arguments for higher budgets come with the promise that fewer officers are necessary if technological tools can reduce crime more broadly. Secondly, another major economic concern is that the analytics industry could capture a large market of policing departments, and the market dynamics at play raise questions about who stands to benefit from the increase in predictive policing as a strategy, tool, and investment.

The impact of technologies designed in military or computer science contexts that are later used for commercial and labor management purposes or are co-opted to criminal justice and policing spaces is unclear and raises a number of questions. Does this mean that data tools are essentially *economic* tools aimed at operational efficiency (the same crime rate with fewer police) or do these tools genuinely produce safer neighborhoods (i.e. *prevent* crime)? Currently, there is limited evidence connecting predictive policing methods to a reduced crime rate, and even reduced crime rates might hide altered or opaque reporting procedures – a difference in reporting, not in policing. Determining the effectiveness of these methods is further complicated because there is no one set of predictive policing methods currently, but a range of formal and informal deployments.

One example of predictive policing that was *not* location based was a program of social network analysis in Chicago. Individuals who were co-arrested with someone who later became the victim of a homicide were themselves found to be of at increased risk of subsequent homicide. However, interventions staged to identify and inform, or identify and not directly inform potential victims have produced unclear results. For example, between districts who told potential victims that they were at high risk compared to districts who did not, there was not a significant effect between the experimental and control group. In addition, there was no consistent policy for acting on the information. Some districts informed the identified individuals of their “status,” others did not. This highlights a central challenge to new predictive policing methods – even when an analysis is performed, what new actions are appropriate to take?

There is additional difficulty generalizing from the Chicago example and other similar cases because these disproportionately monitor and affect low-class communities of color. Not only is the deployment of surveillance programs into these neighborhoods (i.e. Bed-Stuy in Brooklyn) potentially disruptive and intrusive, it also continues the corresponding paucity of data about white, middle-class crime. This

imbalance highlights or magnifies one of the greatest potential downfalls of predictive policing methods – that the opacity of algorithms and the authority of data would function to entrench and reinforce already unjust or negligent policing practices. Ultimately, the workshop came closest to consensus on the idea that the most important matter was what the interventions are in response to risk predictions, as that would make the key difference as to whether citizens (especially persons of color) were better protected or continued to have many negative experiences with the police. These interventions need not be punitive law enforcement activities; they can be provisions of services, protective operations, etc.

Participants also discussed how predictive policing may reinforce past biases. Namely, specific neighborhoods have been subjected to large amounts of highly biased policing (zero tolerance policies/ broken windows policing/stop-and-frisk/etc.) leading to large numbers of “low quality” (aka “garbage”) arrests and convictions, which in turn means that a naïve algorithm might flag those same areas for even more discriminatory policing given the number of “crimes” in the area.

Future Interventions in Predictive Policing

Given that current predictive policing methods are of unclear efficacy at best and potentially harmful at worst, the workshop participants discussed what types of interventions would be most productive in the future. A range of possible changes were described and considered, including: the dismantling of predictive policing programs, the testing of predictive algorithms on other spheres, the reform of the data sources used for prediction, the reconsideration of who acts as a first responder to predictive policing predictions (law enforcement vs. social and community services), and the development of broadly-accepted best practices.

Despite a more general lack of clarity about predictive policing, there have been some documented successes—communities where locations were identified and officers were stationed physically whose crime rates were lower than previous reported years. But what is the avenue for increasing these early successes? Were those early successes were actually successes of predictive policing or related to other environmental factors/crime trends? Should the “black box” of predictive policing be opened, and would better algorithms (SVM, decision trees, neural nets) give *better* crime prevention? Or is it more a matter of data collection, requiring infrastructural and cultural changes to encourage better and more reporting from community members? Community reporting has always been fraught; how does data collection change community members’ understanding of police control over local situations?

Currently, different predictive policing products/vendors (Hunchlab, PredPol) are experimenting with different data analysis techniques. Location identification is still the most common, even if different systems take different inputs into account. However, more work needs to be done to evaluate data inputs and analysis to identify and mitigate bias. And while the ability of certain techniques to identify relevant “hot spots” was debated, the workshop seemed to agree that no matter the algorithm, there needed to be more nuanced responses to data analysis than simply an increase in physical police presence. Most

notably, this might include the communication of risk assessment data to collaborating organizations (not simply police departments), such as social services or civil rights groups.

Areas for Further Exploration

The group agreed that every step of predictive policing was eligible for scrutiny and reform. Better data need to be gathered. Better tools need to be coded. The communities doing the coding need to include more people of color, and more members of affected communities. The methods of predictive policing (both the algorithmic and the *non*-algorithmic) need to be more transparent. Concrete methods of evaluation need to be developed, and the very *purpose* of using predictive analysis in policing needs to be questioned and challenged.