Predictive Policing: Proceed, but with Care

Predictive Policing: Proceed, but with Care

Data-driven analytics can increase the effectiveness and efficiency of police work. Police departments should however proceed with care, as tools such as predictive policing raise a number of concerns regarding human rights and civil liberties.

By Matthias Leese

Data-driven analytical tools such as predictive policing have been hailed as a revolution in police work. Based on algorithmic data analysis, predictive policing software aims to identify crime risks and enable the police to carry out prevention measures in a targeted fashion. This opens up an avenue to render police work more effective and efficient, as better results could arguably be achieved with fewer resources. This potential has been welcomed particularly in light of the overall economic pressures in many countries and the budget cuts that police departments face.

A new study, carried out at the Center for Security Studies and at the University of Hamburg/Technical University Berlin, has investigated how the implementation of predictive policing software transforms police work. Based on multiple years of field research with German and Swiss police departments, the study pays particular attention to the effects that data-driven analytics can have on civil liberties and the relationship between the police and the population.

Data were primarily collected in the form of expert interviews with police officers, crime analysts, and software designers. Moreover, participant observation methods allowed for the observation of crime analysis practices using predictive policing software. Finally, legal, tech-

nical, and operational documentation with regard to predictive policing was assembled. The resulting data were coded and clustered thematically, resulting in the identification of the most prevalent issues and challenges in data-driven police work.

The results of the study have been published in the form of a book, entitled *Criminal Futures: Predictive Policing and Everyday Police Work,* co-authored by Matthias Leese and Simon Egbert (Routledge 2021). The book is available for free download as an Open Access eBook version here. This Policy Perspective summarizes the main

Key Points

▮ Data-driven analytics can help to improve police work and crime prevention.

▮ Tools such as predictive policing can however potentially undermine civil liberties and impair the relationship between the police and the population.

▮ In order to ensure responsible use, policy-makers and police chiefs should critically reflect questions of data, automation, decision-making, communication, and operative measures in algorithmically supported police work.

Chicago Police Officer Grand monitors the Police Observation Devices on computer screens in Chicago. *Joshua Lott / Reuters*

recommendations for a responsible use of data-driven analytics in police work, as presented in the book.

Deploying Data-Driven Analytics Responsibly: 7 Steps

First: Data must not be mistaken for a true representation of the world. They are always a partial account that has been constructed within a particular context and for a particular purpose. A healthy degree of skepticism toward data is appropriate, especially when they are acquired from external sources.

Data do not exist independent of their creation. When police officers produce data from a crime scene, they look for specific things that will allow them to describe their findings. Other observations will be discarded and will not end up as analyzable crime data. The categories that are used to create data are, moreover, already pre-defined by the classification system in the central database. There is thus already a selection bias at work when data about crime and society are created.

Additionally, data creation is prone to error. Evidence at the crime scene might be overlooked, data might be entered sloppily into the database in the late hours of a night shift, or they might accidentally end up in the wrong category.

Another important aspect concerning data is the uncertainty that surrounds crime and criminal investigations. At an early stage, not much might be known about stolen goods, damages, potential suspects, and forensics. This is likely to change throughout an investigation, which is why data need to be amended and updated regularly. Overall, crime data are notorious for their unreliability and incomplete nature, and they need to be subjected to rigid quality control measures.

In summary, any dataset – including data acquired from external sources – contains bias (i.e. the over- and/or underrepresentation of certain phenomena). This is inevitable, but must be kept in mind when evaluating the 'truthfulness' of data and their representative value. As data are used as input for analytical tools such as predictive policing software, there is a danger that data bias will be perpetuated throughout the analysis and live on in the form of biased risk estimates.

Second: Algorithmic crime analysis tools must always remain transparent and comprehensible, independent of whether they are commercial products or in-house developments. Overly complex and/or black-boxed applications will undercut institutional accountability and potentially trigger resistance among police staff.

Algorithms range on a scale from simple and easily understandable to inherently complex and irretraceable even for experts and programmers. Usually, the more complex variants are also the more powerful ones, as they are capable of handling large and heterogeneous datasets or even of 'learning' and adapting to new patterns in the analyzed data.

The inner workings of complex algorithms are often called 'black boxes', meaning that humans can see the data input and the analytical output, but they can no longer understand the processes that took place in between. The likelihood of algorithms becoming black boxes further increases when commercial tools are used, as their design and analytical models are usually considered trade secrets.

For police work, black boxes can have two fundamental implications. On the one hand, black boxes can impinge on the police's capacity to be accountable for their actions towards the public. Accountability hinges on the ability to explain how decisions were made and why specific actions were carried out. When the ways in which data are analyzed are incomprehensible for humans, this ability is essentially lost.

On the other hand, they make internal resistance against data-driven analytics more likely. Police officers are generally skeptical towards new technologies in the first place, and such skepticism can easily turn into rejection when they come under the impression that their own expertise and professional experience becomes overruled by a technological tool that they cannot understand. As a consequence, there is a chance that analytical insights will not be implemented.

Third: Full automation of analytical processes should in principle be ruled out. Human analysts must always remain in the loop and have meaningful control over system functions. That means that algorithmic systems must not withhold information from the user or proceed at critical junctions without user approval.

Predictive policing software and other data-driven analytical tools automate many of the analytical tasks that previously were carried out manually by a human analyst. In this way, intelligence can be produced much quicker, on a larger scale, and without random error. Automation is thus fundamental for the advantages that data-driven analytics bring for police work.

Too much automation, however, has been shown to be detrimental for human control of algorithmic analyses. High levels of automation effectively remove the human from the process and leave little or no possibility for intervention in case of doubt about the data input or about recommendations for action.

Basing police work on data-driven analytics in an unmediated fashion is problematic, not least with regard to bias and accountability. It is thus important to carefully configure automation and human oversight in ways that ensure meaningful control at all times. Only then will police departments be able to benefit from predictive policing and other data-driven analytical tools, while firmly remaining in the driver's seat.

Fourth: Decisions must always be made by humans. In light of possible automation bias, critical engagement with algorithmic recommendations should be encouraged and the right to override them should be facilitated and institutionally enshrined.

Research has demonstrated that humans show a significant tendency to uncritically follow recommendations given by technical systems. The reason for this 'automation bias' is that humans consider technical systems to be objective, neutral, and immune to error.

There are however many potential error sources in data-driven analytics, including data creation, data consolidation, data preparation, theoretical choices, modeling, and the operationalization of variables.

In order to make informed and responsible decisions, human analysts should actively be encouraged to engage with all aspects of the analytical process, including the explicit right to overturn algorithmically produced intelligence and recommendations for action.

Blind trust in algorithmically produced intelligence and recommendations for action might lead to faulty operational decisions that can undercut the effectiveness of police work and deteriorate the relationship between the police and the public.

Fifth: Communication of risk estimates should specifically indicate that risk must not be treated as evidence, but as a possibility. Criminal futures, even when convincingly presented, may not come into being after all.

Data-driven analytics are usually deployed in a preventive fashion, indicating a risk that a certain event might happen in the future if not actively prevented. Based on such risk estimates, police departments can then adjust their operational measures and larger prevention strategies.

There is however a tendency to understand risk estimates as a fact rather than as the possibility that they actually represent. The indicated event – for example a residential burglary in a specific neighborhood during a particular timeframe – might happen, but it might just as well not happen, even if no targeted prevention measures are implemented.

The fact that there is no certainty whether crime risk will materialize should be taken into consideration during the operational planning of crime prevention. Clear communication and an awareness of the characteristics of risk can give stronger attention to balancing targeted prevention measures with complementary randomized and non-targeted measures.

Sixth: Risk estimates can have performative effects on individual behavior. Patrol officers, in particular, should be aware of how their perception might be affected by imaginaries of criminal futures and how this might impact their interactions with citizens.

Patrol officers, when carrying out their tasks on the basis of algorithmically produced intelligence, show a tendency to be more suspicious as compared to randomized patrols. In neighborhoods where there is an allegedly higher risk for

Further Reading

Egbert S. / Leese M., ***Criminal Futures: Predictive Policing and Everyday Police Work*** (London/New York: Routledge, 2021).
This book presents the findings from a multi-year study on predictive policing in Switzerland and Germany. The empirical analysis is embedded in larger theoretical, conceptual, and normative considerations of data-driven tools for police work.

Kaufmann M. / Egbert S. / Leese M., **"Predictive Policing and the Politics of Patterns,"** *British Journal of Criminology* 59:3 (2019), pp. 674−692.
This article advances the theoretical understanding of predictive policing and other data-driven tools for police work by engaging with pattern recognition as the central tenet of algorithmic analysis. It was awarded the 2019 Radzinowicz Prize for the most important contribution to knowledge of criminal justice issues and the development of criminology.

Egbert S., **"Siegeszug der Algorithmen? Predictive Policing im deutschsprachigen Raum,"** *Aus Politik und Zeitgeschichte* 67:32-33 (2017), pp. 17−23.
This article provides a brief overview of predictive policing tools and their use in Germany and Switzerland.

certain types of crime, surveillance and control activities are therefore likely to be intensified.

This can lead to situations where citizens are targeted as suspicious simply by association with their surroundings. When individuals are subjected to controls merely on the basis of allegedly heightened crime risk within the area, this presents a severe violation of the principle of probable cause for police interventions.

Discrimination on the basis of data-driven analytics can also include the aggravation of already problematic police practices that are based on stereotypes, prejudice, or even outright racism. It is thus important to make clear to patrol officers that risk does not imply that crime is in fact bound to happen, and to raise reflexivity in terms of the effects that knowledge of risk estimates can have on their actions.

Seventh: The capacities and limitations of predictive policing must be carefully assessed. It should remain a complementary tool and not replace long-term strategic programs that address the root causes of crime.

On the strategic level, police departments should be careful not to overemphasize the role that data-driven analytics can and should play in police work. In light of political discourse, media attention, as well as financial commitments made through procurement and implementation, there is a perceived need to maximize the utility of predictive policing and other data-driven analytical tools.

Analytics do however have a tendency to favor the prevention of crime rather than to address its root causes. They suppress rather than evaluate why crimes happen and how incentives for criminal behavior could be addressed in the first place. Admittedly, the reasons for the occurrence of crime might often be outside the scope of police work. Nonetheless, it is important that data-driven analytics do not replace programs of community engagement and debates about social reform.

Conclusion

Taken together, the points discussed here present some of the most pressing issues with regard to data-driven analytics in police work. Paying attention to the recommendations will put police organizations in a position to critically assess and reflect how data-driven analytics can be deployed in a responsible fashion.

They apply to current forms of predictive policing, and they will apply equally to future versions of predictive policing and other conceivable analytical tools. In the end, society has a strong interest in neither impeding the capacities of the police as a guarantor of social order nor curtailing civil liberties and human rights.

Dr. Matthias Leese is a senior researcher at the Center for Security Studies (CSS).