Crime prediction for more agile policing in cities - Rio de Janeiro, Brazil

Case study of the U4SSC City Science Application Framework

For more information, please contact:

u4ssc@itu.int

Website: http://itu.int/go/u4SSC
11 SUSTAINABLE CITIES AND COMMUNITIES
Case study: Crime prediction for more agile policing in cities – Rio de Janeiro, Brazil

October 2019
Foreword

This publication has been developed within the framework of the United for Smart Sustainable Cities (U4SSC) initiative.

Acknowledgements

The development of the Case study: Crime prediction for more agile policing in cities has been researched and written by Katherine Aguirre (Igarape Institute), Emile Badran (Igarape Institute) and Robert Muggah (Igarape Institute) and has been edited and revised by Okan Geray (Smart Dubai Office).

The authors wish to thank the U4SSC management team, Nasser Al Marzouqi (U4SSC Chairman) and Victoria Sukenik, Paolo Gemma, Abdurahman M. Al Hassan and Albert Medrán (U4SSC Vice-chairmen) for their respective assistance and contributions. The authors also wish to extend their gratitude to the contributing organizations along with their representatives: Cristina Bueti, Chris Ip and Reyna Ubeda from the International Telecommunication Union (ITU), Paola Deda, Amie Figueiredo and Agata Krause from the United Nations Economic Commission for Europe (UNECE) and Robert Lewis-Lettington from the United Nations Human Settlements Programme (UN-Habitat).

The opinions expressed in this publication are those of the authors and do not necessarily represent the views of their respective organizations or members.

© ITU, UNECE and UN-Habitat.
# CONTENTS

Foreword .......................................................................................................................................................... ii  
Acknowledgements ........................................................................................................................................ ii  
1. Introduction .............................................................................................................................................. 1  
2. The smart project(s) ................................................................................................................................. 2  
3. Conclusions .............................................................................................................................................. 7  
A. References and Bibliography ................................................................................................................ 10  
B. Annex ...................................................................................................................................................... 14  
C. List of discussion partners/interviews .................................................................................................... 15
1. Introduction

1.1. Background

Rio de Janeiro, or simply Rio, is the second-most populous municipality in Brazil and the sixth-most populous in the Americas. The metropolis is anchor to the Rio de Janeiro metropolitan area and is the capital of the state of Rio de Janeiro, Brazil’s third-most populous state. Part of the city has been designated as a World Heritage Site, named "Rio de Janeiro: Carioca Landscapes between the Mountain and the Sea", by UNESCO on 1 July 2012 as a Cultural Landscape.

Rio de Janeiro is headquarters to Brazilian oil, mining, and telecommunications companies, including two of the country’s major corporations – Petrobras and Vale – and Latin America’s largest telemedia conglomerate, Grupo Globo. Being the home of many universities and institutes, Rio is the second-largest center of research and development in Brazil. Rio de Janeiro is also one of the most visited cities in the Southern Hemisphere and is known for its natural settings, including several beaches.

There are significant disparities between the rich and the poor in Rio de Janeiro, and different socioeconomic groups are largely segregated into different neighborhoods. Although the city is ranked as among one of the world’s most populated metropolises, a large number of its inhabitants live in slums known as favelas. There have been a number of government initiatives to counter this problem, from the removal of the population from favelas to housing projects such as Cidade de Deus to the more recent approach of improving conditions in the favelas and bringing them up to par with the rest of the city, as well as the development of the "Favela Bairro" program and deployment of Pacifying Police Units.

1.2. Challenge and response

Crime in Brazil involves an elevated incidence of violent and non-violent crimes. According to a study by the Brazilian Forum of Public Security (a research organization), 63,880 people were murdered in Brazil in 2017, up 3 percent from the year before, resulting in 175 deaths per day. The research indicates that the murder rate in the country was 30.8 per 100,000 people, up from 29.9 in 2016. In comparison to the United States, it had five homicides per 100,000 people in 2015 — down from eight per 100,000 in 1996. Mexico, which is also suffering from a soaring murder rate, had less homicides per capita than Brazil with 25 per 100,000 last year.

Despite the high perception of crime in Brazil, Rio de Janeiro, being one of the most populated cities of the country, has a lower crime rate than Northeast Brazil. However, it is far more criminalized than the southern region of Brazil, which is considered to be the safest region in the country.

Therefore, the challenge for Rio de Janeiro was to predict crime and implement more agile policies to lower crime rate. The response has been to develop a digital platform to predict crimes based on available past crime data and also avail it as an app for residents and visitors of the city.
2. The smart project(s)

Cities are where the future lies. They are hubs of innovation, productivity and experimentation. However, cities also are sites of crime and violence. More than ever, municipal authorities, private firms and civic groups are experimenting with new ways to improve safety in cities. In some cities, new technologies are being deployed to improve the situational awareness of public authorities and citizens. In others, all-encompassing surveillance and monitoring systems are being implemented, raising challenges on the fundamental norms of privacy.

In most developed cities, high-frequency time series information on insecurity is increasingly available. Literally thousands of gigabytes of raw data are available representing the dynamics and characteristics of crime. New high-power computer analysis is giving rise to a next generation of smart, agile and evidence-informed policing strategies. Predictive platforms in particular can enhance police operations, identifying priority targets for police intervention, and enabling more effective allocation of police resources.

2.1. Vision and content

Predictive analytics are hardly new. Statistical and mathematical models have long been used to predict where crime may occur. Predictions are based on a series of assumptions. For example, the criminological literature predicts that violent crime and property crime are not only highly concentrated in specific locations, but also tend to occur at predictable intervals.

Predictive policing tools are being rolled out by police departments across North America, Western Europe and parts of Asia. Police departments typically use thermal maps indicating the locations and times where the probability of crime is highest. Senior law enforcement officers can apply this information to plan their routine operations and send officers and patrol cars to the right locations at the right time.

This so-called "hotspots policing strategy" - merging data analytics with targeted policing – has been around for over two decades. Scientific evaluation studies have shown that it is an effective crime prevention strategy⁵. While concentrating resources on crime hot-spots may contribute to a modicum of crime displacement, it happens less than expected. Indeed, departments must regularly update their data systems and operational strategies as crime itself undergoes structural changes over time⁶.

Predictive policing is also evolving. They are benefiting from advances in machine learning, coupled with more affordable computational power. When compared to traditional hot-spots mapping approaches using retrospective data, predictive analytics can process more granular data at a more rapid pace, generating predictions associated not only to a location, but also to a crime type, and to specific times of the day and days of the week. When applied with fidelity, such tools can help police departments validating their predictions on a daily basis and adapt their responses accordingly to the everchanging crime patterns.
Crime forecasting models are another example of utilizing new technologies for more agile security: promising data-driven and problem-oriented approaches that can speed up decision-making and providing smart solutions that help reduce human biases and inefficiencies. In an era marked by the fourth industrial revolution, AI, the Internet of Things and Big Data are available to help law enforcement and criminal justice authorities adopt more effective policing strategies.

In 2016, the Igarapé Institute partnered with Via Science - a data analytics firm - to develop the CrimeRadar app, a public-facing crime forecasting platform that evaluates relative crime frequencies in different locations and times in the metropolitan of Rio de Janeiro.

### 2.2. Implementation

CrimeRadar is a digital platform that forecasts the probability of crime. It runs on smartphones and desktop browsers. The software uses advanced data analytics to show real and relative crime rates and risks for different neighborhoods at different times in the Rio de Janeiro municipality.

![Figure 1: Illustrative CrimeRadar Platform](image)

CrimeRadar visualizes the safety levels in specific locations and times. By making crime data more accessible and transparent, it improves security for citizens.

The underlying crime data was retrieved from the state Institute for Public Safety and official crime records produced by the state civil police. The platform was launched during the Rio Olympics in August of 2016. Residents and tourists could access the website to view the predictions displayed on an intuitive mobile heatmap.

**Solution Development:** CrimeRadar was conceived by the Igarapé Institute and developed with partners Via Science and Mosaico. The Igarapé Institute provided expert knowledge about the region and worked with various data providers to gather and verify the accuracy of the historical data.
Meanwhile, Via Science created the mathematical engine behind the app by developing algorithms using their advanced machine learning and proprietary software architecture. The Igarapé Institute worked with Via Science to test the results.

Finally, Mosaico, a local software firm, helped design the mobile app interface for the Via Science algorithms and build out an immersive user experience. Mosaico and Igarapé Institute also designed the historical data platform.

**Figure 2: Illustrations of the CrimeRadar Mobile App**

**Why CrimeRadar?**: CrimeRadar is an example of a digital platform to improve personal safety and security. The goal is to improve transparency and access to information for citizens and visitors. CrimeRadar makes crime data accessible, interactive and useful.

Until recently, basic crime data in Rio de Janeiro was inaccessible to the wider public. Those data were often poorly collected and registered. However, there has been a major shift in data management. The newly launched ISPGeo, developed in cooperation with ISP, the Igarapé Institute and several city partners, provided crime data to the wider public in a safer and more accessible manner.

**CrimeRadar Difference**: Other applications that predict crime activity tend to be restricted to police departments and not made available to the public. It is often restricted to a small number of users and the underlying methods are not transparent. By contrast, CrimeRadar is a step towards empowering the wider public to get involved in crime prevention. It is part of a wider movement to make data more open and accessible.
CrimeRadar applies leading edge mathematics, advanced machine learning tools, and an easy to use interface to help translate historical crime data into more accessible and actionable information for users. CrimeRadar is not intended to prescribe or mandate safety levels in a region. Rather, it is intended to serve as one additional data point that the public can use when making decisions about traveling to specific locations.

CrimeRadar is an example of public private partnership whereby actual city crime data was provided to Igarape Institute and its partners to develop a solution which helps both city police as well as citizens and visitors of the city. The Institute provided expert knowledge about the region and worked with various data providers to gather and verify accurate historical data. Via Science helped to create and test the formula behind the app using their proprietary machine learning software architecture. Finally, Mosaico, a local software firm, helped design the mobile app interface.

### 2.3. Results

Given its initial success in Rio de Janeiro, starting in 2018, the Institute partnered with the state military police of Santa Catarina to develop and pilot a police-facing version of CrimeRadar. In 2020, the initiative will be evaluated using a randomized controlled trial (RCT) to assess the effectiveness of predictive policing on the planning of police patrol itineraries and scheduled operations. The intention is to assess changes in crime levels and crime displacement as well as average police response times and public trust in the police.

More importantly, the entire development process and resulting crime forecasting algorithm are being documented in a "social impact statement". The intention is to describe the challenges associated with (and steps taken to ensure) the monitoring and analysis of crime data. The Igarapé Institute is committed to designing and implementing crime forecasting algorithms in a publicly accountable manner.

To start, the software license will require that all police departments deploying the predictive tool comply with a minimum set of transparency and reporting standards. A list of five minimum requirements are set out below and were prepared by the FAT/ML\(^8\) work group, a community of researchers and practitioners concerned with fairness, accountability, and transparency in machine learning:

- **Responsibility and Recourse** – Make available externally visible avenues of redress for adverse individual or societal effects of an algorithmic prediction system, and designate an internal role for the person who is responsible for the timely remedy of such issues.

- **Explainability** – Ensure that algorithmic predictions as well as any data driving those predictions can be explained to end-users and other stakeholders in non-technical terms.

- **Accuracy** – Identify, log, and articulate sources of error and uncertainty throughout the data sources so that expected and worst case implications can be understood and inform mitigation procedures.
• **Fairness** – Ensure that algorithmic predictions do not create discriminatory or unjust impacts when comparing across different demographics.

• **Auditability** – Enable interested third parties to probe, understand, and review the behavior of the algorithm through disclosure of information that enables monitoring, checking, or criticism, including through provision of detailed documentation, technically suitable APIs, and permissive terms of use.

CrimeRadar was featured in the following prominent publications (see figure 3):

**Figure 3: CrimeRadar featured in several prominent publications**

![CrimeRadar featured in several prominent publications](image)

CrimeRadar has had the following impacts:

• It has helped Rio de Janeiro residents and visitors by providing a simple visual heatmap tool which displays various city locations and their likelihood for crime.

• It has incorporated new crime data into Rio de Janeiro’s public security system as it becomes available.

• The city residents’ quality of life has been enhanced by better safety and security.

• Increased safety has increased economic activities in Rio de Janeiro and will likely to continue to be the case in the long run.

• It has helped in deploying police force more effectively to avoid crime by predicting it before it happens.

• It has created operational efficiencies for the police force by reducing and optimizing various costs such as patrolling, focused deployment of police force, etc.

• Reduced crime rate has improved the image of Rio de Janeiro as a highly visited global city.

• Reduced patrolling by police vehicles has also reduced CO2 emissions.
Globally, there is still comparatively mixed evidence of the accuracy of crime prediction, its impact on clearance rates, whether it improves response times or even leads to significant reductions in crime. The only way to really gauge the impacts of crime forecasting is to conduct statistical evaluations that isolate the effects of the measure, including randomized control trials (RCTs). The annex provides a summary of the findings from a number of the most well-known instances of crime prediction implementation.

3. Conclusions

CrimeRadar is an innovative AI-based solution that utilizes historical city crime data to predict crime. The solution is adaptive in that it is able to incorporate new incoming data and make adjustment accordingly. The prediction algorithm can be applied to other cities as long as the crime data is available for them, hence CrimeRadar is by design transferable to other cities (subject to data consistency and translation where needed).

Limitations of crime prediction: For all their promise, predictive crime analytics are not a panacea. For one, certain types of crime—including domestic and interpersonal violence—is not easily amenable to predictive models since they are seldom concentrated in specific locations, and cannot be readily attributed to specific profiles of victims. While predictive algorithms may reduce certain forms of human bias by reducing subjectivities, they ultimately rely on often flawed crime data with systematic reporting biases.

Furthermore, predictive policing experiences, when subjected to closer scrutiny, have registered a host of challenges. They are often costly owing to data storage, lacking transparency in relation to the underlying algorithm, and having on occasion led to the violation of basic rights and civil liberties. Without high quality data and due care in the way they are built, predictive algorithms can unintentionally reproduce and exacerbate societal prejudices. Identifying biases in data sets is complex, requiring deep knowledge in statistics, mathematics, and programming. As with most policing technologies, successful application requires a comprehensive approach. It depends not just on institutional leadership and the technical capacity of law enforcement agencies to incorporate predictive tools into routine operations, but also the development of minimum standards for responsible development, auditing and evaluation.

Designing and deploying predictive tools: Crime forecasting is by definition a mixed method, involving a host of integrated tasks. These include time-series modeling, intensive data mining, hot-spots analysis, and socio-temporal assessment applied to historical crime data. It is important to stress that predictive policing goes beyond basic online mapping tools that track crime.

Statistical methodologies include the near-repeat theory and crime hot-spot analysis. These approaches assume that once a particular violent or property crime occurs in a particular location, it is likely to occur again in that same area. Meanwhile, the risk terrain model is more focused on geographical analysis, seeking to identify risk factors and features of crime-affected locations, such as insufficient public lighting and potential escape routes.
Predictive policing is grounded on several established theories of crime behavior and crime opportunity explaining crime concentration and repetition, and why crime occurs in some places and not in others. These include:

- **Routine Activity Theory** that states that crime depends on multiple factors including the motivation of offenders, suitable targets and an absence of capable guardians\(^1\);

- **Rational Choice Theory** underlines that criminals make rational decisions based on opportunity and estimated costs such as the possibility of being imprisoned and punished\(^2\); and

- **Crime Pattern Theory** that explains why, when and where crime happens, focusing on the intersections and commonalities between victims and perpetrators\(^3\).

By applying different combinations of these methods and theories, a growing number of universities and commercial vendors have developed predictive policing software, serving to some of the world’s largest police departments. Existing commercial solutions can be broadly categorized into two categories: (i) methods that predict the location of crime; and (ii) methods that predict likely crime offenders and victims.

The first method involves processing historical police reports, emergency hotline calls, weather forecasts and even the locations and dates of large public events to calculate the probability of crime happening in space and time. The second - and perhaps more controversial - method often processes arrest data\(^4\) including criminal records and social media profiles as well as location, race, age, gender, and ethnic data. The method generates a shortlist of high risk individuals who are determined to be potentially involved in future crime.

**Controversies associated with crime prediction:** There are widespread concerns that predictive policing tools could unintentionally exacerbate over-policing of marginal areas and undermine privacy. It is widely known that algorithms can reproduce existing patterns of discrimination, reinforcing previous errors and biases of programmers and embedded in databases. There are very real ethical questions about the extent to which such tools can influence police to disproportionately surveil marginalized neighborhoods and communities. Related, there are fears that such tools may augment race and age profiling and undermine privacy rights and civil liberties.

Recent studies funded by the US National Science Foundation demonstrate how predictive policing models are susceptible to ‘runaway feedback loops’\(^5\). In these cases, police are repeatedly sent to the same identified hot spots, irrespective of the true crime rates. The researchers demonstrated how historical crime incidents that have been "discovered" by on-duty officers can aggravate the degree of runaway feedback, while in turn, historical incidents that were "reported" by citizens can attenuate, but cannot entirely remove such feedback.

The accelerated pace and spread of crime and violence prediction tools means that these concerns will only grow in the coming years. Indeed, new platforms are already being tested that aim to automatically classify gang-related crime\(^6\), combine social media with criminal history to predict crime\(^7\), and use artificial intelligence to identify individuals\(^8\) with higher risk profiles of committing
terrorist acts. The rapid roll-out of these tools invariably raise complex ethical questions in relation to police action and civil rights.

**Establishing standards and regulations:** There are several considerations that law enforcement agencies would do well to consider before implementing predictive policing systems. First and foremost, departments must evaluate the quality of their crime data and the capabilities of officers and officials to perform such an evaluation. Specifically, it is imperative that crime underreporting and blind-spots are dealt with, and that departments can ensure that citizens from all neighborhoods and social groups have confidence—and make use of—the police emergency hotline to report crime when in need. Likewise, police reports must indicate precise addresses including their geographic coordinates. Also, to help mitigate the risk of runaway feedback loops, incidents must be labeled to indicate if they were reported by citizens, or if they were initiated by an on-duty officer while on routine patrol.

Capital and operational expenditures associated with crime prediction tools must also be carefully assessed. The most sophisticated forecasting packages are expensive and may not be suitable for all police departments, especially smaller and mid-sized units in low- and medium-income settings. Instead of purchasing expensive software, some police departments may benefit more from hiring and training analysts to use standard (and often open source) software to plot crime events on a map and run simple (yet useful) time series analysis.

Policing innovations for agile security also should make use of the interconnection of urban infrastructure including sensors and unstructured data. Even so, privacy concerns should be paramount in the decision to process such information. Where possible, predictive tools should allow citizens to understand what is inside the “black box”. While private vendors understandably seek to protect their source code, this lack of transparency (coupled with their underlying mathematical complexity) makes it difficult for law enforcement agencies and civil society to understand how the predictions are generated. This can undermine confidence in the tool. Complicating matters, the secrecy associated with predictive tools may subject departments to increasing legal liabilities as cybersecurity and privacy regulations continue to evolve.

Above all, people must remain the most important element in the crime forecasting process, even when the most advanced software packages are used. Predictive tools need not replace the intuition and experience of law enforcement officers, but rather complement them in a transparent and auditable manner. When responsibly implemented, predictive policing tools can improve law enforcement’s capabilities to solve problems, make decisions, and more effectively plan their operations.
A. References and Bibliography


## B. Annex

**Selected applications of predictive policing and evaluation results**

<table>
<thead>
<tr>
<th>City, date of the evaluation, name of software</th>
<th>Finding</th>
</tr>
</thead>
<tbody>
<tr>
<td>Shreveport, Louisiana (US), 2012, Predictive Intelligence Led Operational Targeting (PILOT)</td>
<td>Formal evaluation (blocked randomized controlled field experiment). No statistical evidence that crime was reduced more in the experimental districts than in the control districts.</td>
</tr>
<tr>
<td>Los Angeles (US), 2013, PredPol</td>
<td>Formal evaluation (randomized control trial). Average 7.4% reduction in crime volume as a function of patrol time. Reduction of property crimes by 12% compared with the previous year in treated area (Foothill); in neighboring districts, property crime rose 0.5%. Note: non-independent evaluation, done by founders of PredPol.</td>
</tr>
<tr>
<td>Chicago (US), 2013, Strategic Subjects Litc- SSL</td>
<td>Quasi experimental evaluation. No impact of the list of people most likely to be involved in a shooting</td>
</tr>
<tr>
<td>Greater London Area (UK), 2013, Metropolitan Police Service (MPS) algorithm (‘M8R’)</td>
<td>Evaluation of crime forecasting accuracy. Burglary – ‘very low’ to ‘low’ predictive accuracy (hit rates of 0 – 5%). Theft from motor vehicle – ‘low’ predictive accuracy (hit rates of 1-10%). Robbery – ‘low’ to ‘medium’ predictive accuracy (hit rates of 0-20%). Theft from person – ‘medium’ to ‘good’ predictive accuracy (hit rates of 13 - 54%).</td>
</tr>
<tr>
<td>Kent (UK), 2014, PredPol</td>
<td>Operational review. PredPol is 10 times more likely to predict the location of crime than random patrolling and more than twice as likely to predict crime as boxes produced using intelligence led techniques. During the North Kent pilot 25% of boxes were visited on average and a 4% reduction in crime was observed.</td>
</tr>
<tr>
<td>Milan (IT), 2008-17, KeyCrime</td>
<td>Quasi random evaluation. Increase in clearance rates. Reduction of robberies in 18%. Saving in prevention of violence up to USD$2.5 million.</td>
</tr>
<tr>
<td>Richmond, Virginia (US), 2006-17, WebFOCUS -IBM SPSS’s Clementine and Predictive Enterprise Services</td>
<td>Report of results (no evaluation). Since implementation, reduction of incident rates of murder (32%), rape (20%), robbery (3%), aggravated assault (18%), burglary (18%) and auto theft (13%).</td>
</tr>
<tr>
<td>Santiago de Chile (CL), 2015, CEAMOS</td>
<td>Report of results (no evaluation) 89% of effectiveness in the tests carried out by the police</td>
</tr>
<tr>
<td>La Plata (AR), 2018</td>
<td>Report of results (no evaluation). Reduction of crime in 40% in identified hot spots.</td>
</tr>
<tr>
<td>Durham (UK), 2013, Harm Assessment Risk Tool (HART)</td>
<td>Royal United Services Institute study. HART was found to predict low-risk individuals with 98 percent accuracy and high-risk with 88 percent accuracy.</td>
</tr>
<tr>
<td>The Netherlands, 2017, Crime Anticipation System (CAS)</td>
<td>Trial pilot. Over 30% of thefts were committed in the zones predicted by the algorithm.</td>
</tr>
<tr>
<td>Baden-Württemberg (GE), 2016, PRECOBS</td>
<td>Max Planck Institute evaluation. Moderate effects in the reduction of burglary</td>
</tr>
</tbody>
</table>
C. List of discussion partners/interviews

This case study has been prepared by:

Katherine Aguirre  
Senior Researcher, Igarape Institute

Emile Badran  
Senior Project Advisor, Igarape Institute

Robert Muggah  
Research director and program coordinator for citizen security, Igarape Institute

This case study has been edited and revised for U4SSC compliance by:

Okan Geray  
Strategic Planning Advisor, Smart Dubai Office
Endnotes

1 Section 1.1 is based on Wikipedia data on Rio de Janeiro.
5 Braga, Papachristos, Hureau, 2012.
7 Muggah, 2018.
9 Knight, 2017.
10 Lum, Isaac, 2016.
11 Oram, 2016.
14 Cullen, Wilcox, (n.d.).
15 Lum, Isaac, 2016.
16 Ensign, Friedler, Neville, Scheidegger, Venkatasubramanian, 2017.
20 Olligschlaeger, (n.d.).
For more information, please contact:

u4ssc@itu.int

Website: http://itu.int/go/u4SSC