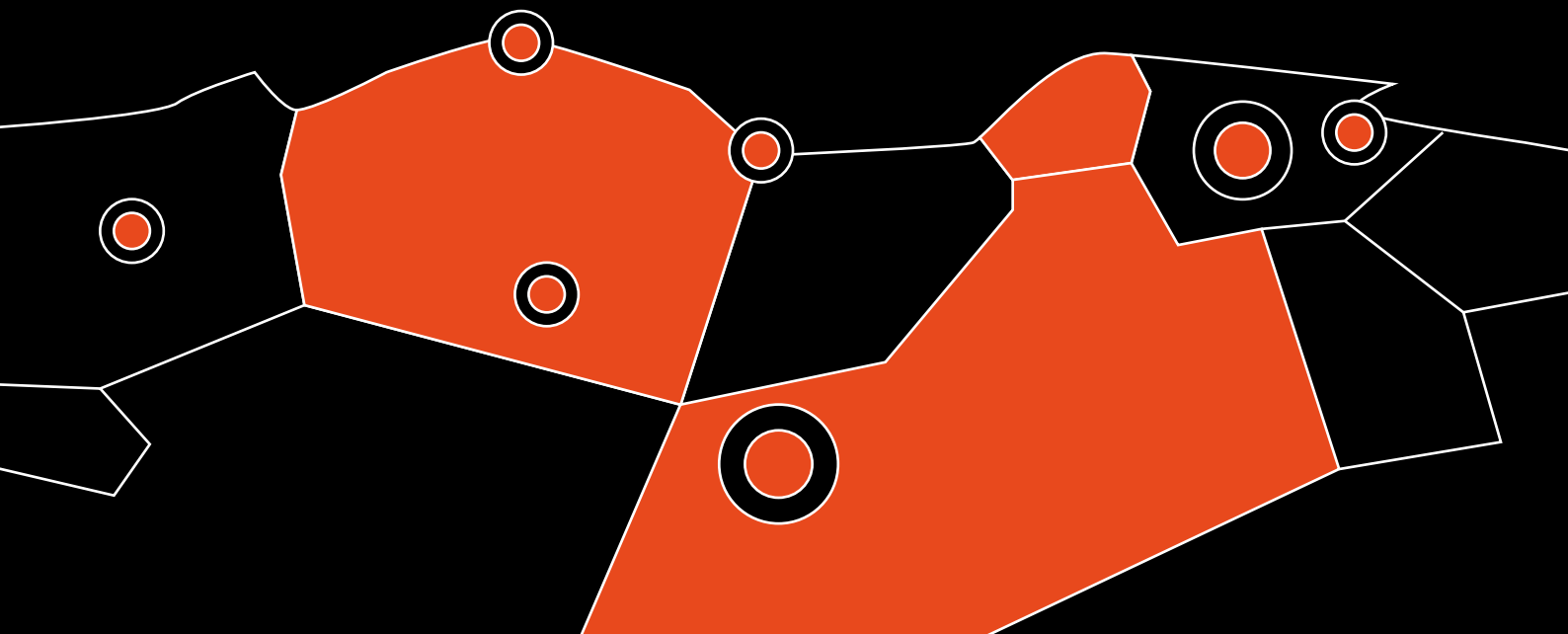
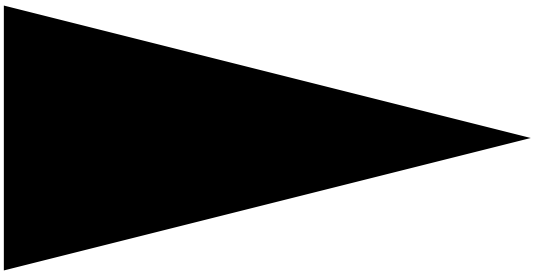
La Quadrature du Net

# ‘Predictive’ Policing in France:

## Against Opacity and Discrimination, Why a Ban is Needed

January 2025





# Summary

This report focuses on police use of ‘predictive’ policing technologies in France. These data-based algorithmic and automated computer systems seek to analyse and ‘predict’ where crime will occur in certain locations. Police and enforcement action can then be taken in those areas. This is also known as ‘hotspots’ or ‘hotspot policing’.

After describing the aim of this research and our method of investigation, we provide some background information on the institutional structure of the French police and their databases.

The report then offers a summary of the more recent developments on the subject, focusing particularly on three ‘predictive’ policing technologies:

- PAVED, a tool developed by the Gendarmerie Nationale;
- Smart Police, a product sold by startup Edicia to municipal police forces; and
- M-Pulse, a project formerly known as the Observatoire Big Data de la Tranquillité Publique (the Big Data Observatory of Public Safety), developed by the city of Marseille in partnership with Engie Solutions.

This is followed by analytical section. We highlight that these systems apparently fail to achieve their objectives, with little to no impact on crime. However, they result in police abuses of power and the increasing surveillance and targeting of communities already impacted by structural discrimination.

These concerns are especially pronounced given the lack of evaluation, oversight and transparency surrounding these algorithmic systems. We are therefore calling for a ban on ‘predictive’ policing systems.

# Table des matières


Summary 3

Introduction 7

Methodology 8

Outline and main findings 9

1. Technical and institutional context 10

1.1 Organisation and jurisdiction of law enforcement agencies in France 10

1.2 Different approaches to technology 13

1.3 A few observations on police databases 13

2. Overview of 'predictive' policing systems 21

2.1 The origins of criminological mapping 21

2.2 Using data and statistics to direct police action 24

2.3 'Predictive' policing systems 25

3. Critical analysis of these systems 58

3.1 Potentially discriminatory parameters 58

3.2 Feedback loops 59

3.3 False criminological assumptions 60

3.4 Possible abuses of power 61

3.5 Correlation is not causation 62

3.6 The limited effectiveness of this technology 63

3.7 Data collection and use: serious shortcomings 64

Conclusion: prohibit 'predictive' policing 67

# Introduction

This report aims to provide an overview the 'predictive' policing systems used by the French police. These are automated decision-making (ADM) systems designed to focus police activity towards certain areas that are 'predicted' or otherwise deemed more prone to crime and other public order disturbances.

The report is part of the Technopolice action research campaign, led by the collective La Quadrature du Net.[1] The campaign seeks to:

- document the use of new surveillance technologies in urban public spaces;
- identify the political issues they raise; and
- combat the infringements on public liberties they present.

This report is one of a series of reports on automated policing technologies used in Europe, coordinated by the British NGO Statewatch.

We documented the emergence of 'predictive' systems used by the police and advocated for a ban on them as early as 2017.[2] We were then confronted with a lack of up-to-date information and genuine public debate on these systems. As a result, we decided to investigate in more detail. For this report, we gathered available information on several 'predictive' policing systems formerly or currently used by French police forces, including:

- **RTM (Risk Terrain Modelling):** a 'situational crime prevention' system used by the Paris Police Prefecture to target areas for intervention based on 'environmental' data (such as the presence of schools, shops, metro stations, etc.);
- **PredVol**: a system developed in 2015 by Etalab (a French public administration that aims to improve public service and public action through data) to assess the risk of car thefts, tested in the Val d'Oise region in 2016 but scrapped by 2018;
- **PAVED**: a system developed by the *Gendarmerie* from 2017 onwards and trialled from 2018 in various metropolitan departments to assess the risk of car thefts or burglaries. In 2019, shortly before its planned nationwide roll-out, the project was put 'on hold';

1. Visit the Technopolice campaign website at: https://technopolice.fr.
2. La Quadrature du Net — Technopolice, "La police prédictive progresse en France. Exigeons son interdiction !", Technopolice (blog), July 23, 2020, available at: https://technopolice.fr/blog/la-police-predictive-progresse-en-france-exigeons-son-interdiction.

- **M-Pulse:**[3] a system developed by the city of Marseille in partnership with the company Engie to assess the suitability of municipal police deployments in the city's public spaces;
- **Smart Police:** a system that includes a 'predictive' module developed by French startup Edicia which, according to its website, has been sold to more than 350 municipal forces.

## Methodology

This report draws on information from several sources. Firstly, CADA requests (requests for access to administrative documents). These allow members of the public to request various types of documents from institutions and local authorities. Submitting these requests is a straightforward procedure that involves sending an e-mail or letter to an authority specifying the nature of the requested documents.[4]

At the beginning of our investigation in 2022, we submitted several rounds of CADA requests to the Ministry of the Interior, local authorities, and the Commission Nationale de l'Informatique et des Libertés (CNIL, or National Commission on Informatics and Liberty). CNIL is an independent French administrative regulatory body tasked with ensuring that data privacy law is applied to the collection, storage, and use of personal data.

We requested information on software, impact studies, internal documentation, correspondence between companies and local authorities, and even the source code of the system. Unfortunately, very few of these requests were successful. It is worth noting that France performs very poorly in the 'freedom of information' index.[5]

We also conducted open-source research:

- reviewing publicly-available documents about the use of these systems;
- investigating websites and documentation of the companies producing these technologies;
- and listening to media appearances of their promoters and manufacturers.

3. Previously named Observatoire Big Data de la Tranquillité Publique (the Big Data Observatory of Public Safety).
4. See our article detailing the CADA request procedure: *https://technopolice.fr/blog/guide-se-renseigner-sur-la-surveillance-dans-sa-ville/*
5. European Public Accountability Mechanisms,"France Public Accountability Index", 2022, available at: *https://europam.eu/?module=country-profile&country=France#info_FOI*; Global Right to Information Rating (blog), "France Country Details," September 2011, available at: *https://www.rti-rating.org/country-detail/*

Although academic research on 'predictive' policing systems used in France is limited, we reviewed the available scientific articles that have been published. We also held discussions with researchers working on the subject, people who have worked in the *Gendarmerie* or police services on a given system, and other sources who helped deepen our understanding of the subject.

Ultimately, as with other topics covered by the Technopolice campaign, the main obstacle to our investigation lies in the institutional opacity surrounding these 'predictive' policing technologies and their uses. The other investigative researchers we met with encountered the same difficulties. Aside from a few publications and rare comments by their creators in the press, the level of transparency surrounding 'predictive' policing systems is completely unacceptable considering the serious technical, legal and political issues they raise.

## Outline and main findings

In the first part of the report, we provide a brief overview of the organisation of police forces in France and some of the common uses of data-based technologies in the field of policing, such as databases and facial recognition.

Secondly, we present various 'predictive' policing systems used in France. There are four detailed case studies on Predvol, PAVED, Smart Police and M-Pulse. We also cover two other systems: the Risk Terrain Modeling (RTM) algorithm used by the Paris police force, and MapRevelation, a system supplied by a start-up to national and municipal police forces.

In the third section, we present a critical analysis of these systems.

The main conclusion of this report is that these systems have been developed, implemented and tested for many years in France, with almost total opacity. Despite this lack of transparency, these systems rely on biased data and dangerous criminological theories, and threaten civil liberties and human rights.

Indeed, the available data tends to show that these automated and data-based systems contribute to the amplification of police surveillance and oppression of people and groups already subject to systemic discrimination, by justifying both increased surveillance and police controls against them.

This is why we are calling for an outright ban on 'predictive' policing systems.

# 1. Technical and institutional context

## 1.1 Organisation and jurisdiction of law enforcement agencies in France

### 1.1.1 At the national level

France's *Police Nationale* was created under the Vichy regime in 1941. The *Police Nationale* reports to the Ministry of the Interior, and their officers are civil servants. The *Gendarmerie Nationale*, on the other hand, was historically part of the Ministry of Defence, and since 2009 has been under the authority of the Ministry of the Interior. Gendarmes still have military status and are military personnel.

In practice, the *Police Nationale* and the *Gendarmerie* do not operate in the same geographical areas. The *Police Nationale* tends to operate in urban areas, while the *Gendarmerie* is responsible for rural regions. The distinction between the *Police Nationale* and the *Gendarmerie* is also reflected in the way they organise their work. They have different hierarchical structures, facilities, equipment and technologies.

However, the two institutions have similar prerogatives in terms of policing powers. They both hold judicial police powers (the ability to investigate criminal offences) and administrative police powers (the prevention of public order disturbances, including public order, safety, and health). Administrative policing is considered 'preventative', in that it seeks to avoid the occurrence of offences and disturbances to public order. Judicial policing, on the other hand, has repressive powers. Under the authority of a judge, its task is to record criminal offences, identify the perpetrators and gather evidence.

The distinction between judicial and administrative policing can have significant legal and operational consequences. For example, an identity check can be carried out on the basis of either criminal or administrative law.

Identity checks carried out as part of a criminal investigation involve:

- identifying a person suspected of having allegedly committed or attempted to commit an offence;
- identifying a person who is about to commit a crime or offence; or
- identifying a person who is simply likely to provide information useful to the investigation of a crime or offence.

Administrative identity checks are carried out in the absence of any suspicion of an offence, for reasons relating to the prevention of public order disturbances.

In France, identity checks (whether for a criminal investigation or administrative reasons) must meet certain criteria justifying their necessity. Certain guarantees are supposed to ensure that they do not infringe on civil liberties and human rights. For example, administrative identity checks require objective evidence of a risk that public order will be undermined, in particular the safety of persons or property.[6]

### 1.1.2 At the municipal level

In addition to national police forces, France also has a local level of policing: the *Police Municipale*. It differs from the national police force in a number of ways, not least by the fact that a municipal police force is created by the mayor's office, and can only operate within the municipality's territory.

The *Police Municipale* take their orders from the mayor, but report to the Prefect of Police, who in turn reports to the Ministry of the Interior. Often, an agreement between the municipal and national police forces defines the nature and location of their respective interventions, as well as the way in which they collaborate.

6. Ministry of the Interior, "Mémento policiers municipaux et gardes champêtres", November 10, 2021, p.19. Available at: *https://mobile.interieur.gouv.fr/Media/MI/Files/Memento-Policiers-municipaux-et-gardes-champetres*

The following section of this report illustrates how these agreements between municipalities and the Ministry of the Interior can lead to them sharing specific data-sharing equipment or infrastructure (e.g. video surveillance systems).

The administrative policing powers of the *Police Municipale* are limited. They essentially include:

- a public health mission (epidemic prevention, waste collection, etc.);
- security of residents and their property (surveillance of public buildings, schools, etc.) and public spaces (collaboration with emergency services in the event of accidents or natural disasters);
- maintaining order on the streets during gatherings or demonstrations; and
- public order.[7]

The law confers limited judicial powers on municipal police officers, allowing them to record criminal offences and issue fines through notices of offence.[8]

However, municipal police officers have far more extensive powers than the penal code would suggest. They can arrest a person in the act of committing a felony or misdemeanour punishable with imprisonment, and take them to the nearest *Police Nationale* or *Gendarmerie* station. They can also draw up reports and notices concerning any felony, misdemeanour or offence witnessed by municipal officers. These can either be passed on to the *Police Nationale* or *Gendarmerie*, or to the mayor. The mayor, acting as a judicial police officer, then passes the information on to the public prosecutor.

The law does not authorise the *Police Municipale* to carry out identity checks. However, municipal police officers can gather information about a person's identity, as long as they do not ask for proof of identity. They are authorised to ask for proof of identity in the case of certain offences that fall within their remit.

7. Ibid.
8. They can, for example:= issue fines to owners of dangerous dogs who fail to comply with the law; visually inspect and search (with the owner's authorisation) bags and luggage at public events or at the entrance to a municipal building; issue fines for disorderly conduct such as littering, disturbing the peace at night, or letting dangerous animals roam freely; and record most traffic offences committed on municipal territory when they do not require an investigation.

## 1.2 Different approaches to technology

The *Police Nationale*, the *Gendarmerie Nationale* and municipal police forces all have different approaches to technology. Unlike the *Gendarmerie*, the *Police Nationale* is more accustomed to outsourcing the production of data-based systems, in particular by purchasing them from companies. Attempts to pool the technical architectures used by the two institutions have regularly led to infighting.[9]

As far as municipal police forces are concerned, they are generally equipped with technology by the local authorities of the cities they are part of through bids and contracts with private companies. They do not develop their own technologies. However, as we discuss later, they can sometimes make agreements with the *Police Nationale* or the *Gendarmerie* to share certain equipment or infrastructure, such as video surveillance or the automated decision-making systems mentioned in this report.

## 1.3 A few observations on police databases

In this section, we discuss the main police database used in France, the TAJ database. We go on to examine other important aspects of the French police forces' information-gathering practices.

### 1.3.1 The TAJ database

The *Traitement des Antécédents Judiciaires* (TAJ, or 'Treatment of Criminal Records') database is the main police database used in France for the day-to-day activities of law enforcement agencies. Created in 2012,[10] it merged two previously separate databases used by the *Police Nationale* (STIC) and the *Gendarmerie Nationale* (JUDEX).[11]

9. See for example: Antoine Albertini, "Les occasions perdues d'un logiciel partagé par la police et la gendarmerie," Le Monde, November 30, 2021. Available at: *https://www.lemonde.fr/societe/article/2021/11/30/les-occasions-perdues-d-un-logiciel-partage-par-la-police-et-la-gendarmerie_6104135_3224.html*
10. Covered by articles 230-6 et seq. and R40-23 et seq. of the Code of Criminal Procedure.
11. See detailed information on the TAJ database in the following report: Caisse de Solidarité de Lyon, "La folle volonté de contrôler tout : 50 fichiers de police-justice et les moyens d'en sortir" 2021, p. 29. Available at : *https://iaata.info/IMG/pdf/fichiers_police_justice_renseignement_v3_novembre_2021.cleaned.pdf*

The TAJ database includes information about people suspected of having committed a serious offence as perpetrators or accomplices, as well as victims of offences. A wide range of data is recorded:

- identity;
- occupation;
- family status;
- nationality;
- postal address;
- telephone number;
- e-mail address;
- photographs;
- alleged offences;
- physical characteristics (including body marks and tattoos);
- dates of alleged offences; and
- other data and images relating to offences.

A software called GASPARD supplies this data to both the TAJ database and the *Fichier Automatisé des Empreintes Digitales* (FAED, or Automated Fingerprint Database).

The TAJ database can be updated by the main national police forces (*Police Nationale*, *Gendarmerie Nationale*, customs, intelligence services, tax authorities, and public prosecutors). However, municipal police forces do not have access to it, even for basic checks. During a criminal investigation, authorised agents from the *Police Nationale*, *Gendarmerie*, customs, judicial services or magistrates may consult the TAJ database.

The database can be consulted during administrative investigations to conduct background checks for recruiting people for certain professions (police officers, security guards, military personnel, and workers in the nuclear sector). Police may also use an individual's entry in the TAJ database to inform the media that a person murdered by drug traffickers or a victim of police violence was 'unfavourably known' to the police.

## 1.3.2 Facial recognition practices

Since at least 2011,[12] the Ministry of the Interior has been able to cross-reference the database underlying TAJ with 'facial comparison' software. This enables automated cross-referencing between a suspect's image (collected, for example, using a video surveillance system or from social media) and the photographs contained in the database. The tool used until at least 2020 was Cognitec's FaceVACS-DBScan.

12. Jean-Marc Manach, "Deux millions de 'contrôles au faciès'", OWNI, June 1, 2012. Available at : *https://web.archive.org/web/20210807071515/http://owni.fr/2012/06/01/deux-millions-de-controles-de-facies/index.html*

The TAJ database now contains a total of almost 20 million individual records, and around 10 million facial photographs.[13] The system is used on average more than 1,600 times a day by the *Police Nationale* and the *Gendarmerie*.[14]

Facial recognition is currently only legally authorised for use after an event, in the context of judicial investigations, and using data from the TAJ database. However, evidence shows that administrative identity checks using the facial comparison feature have increased over the past few years—practices which amount to database misuse.[15]

This process starts with a photo of the person taken by police with an official or personal tablet or phone, which is then sent to a colleague with access to a computer. Next, the TAJ database is queried, providing a wealth of information. If the person is in the database, this means that they have already had dealings with the police, and are therefore known to them. Such a 'match' enables the police to check that the person has indeed given their correct identity.

As an example, the use of the TAJ database for administrative identity checks was covered in a news broadcast in January 2020. In this report, a *Police Nationale* officer carries out a biometric comparison during an identity check.[16] A testimony given during an operation in Montreuil in February 2021 also mentions the use of facial recognition to identify people who have been stopped and interrogated.[17]

Political authorities are currently reluctant to legalise these practices more widely, due to the potential political risks it would entail. However, the technical infrastructure enabling real-time use of facial recognition is already in place.

First, there are 90,000 video surveillance cameras in streets throughout France, which are all points of collection for facial images. Second, there are centralised databases that link identity photos to civil status data.

13. Didier Paris and Pierre Morel-À-L'Huissier, "Rapport sur les fichiers mis à la disposition des forces de sécurité", Commission des Lois constitutionnelles, de la législation et de l'administration générale de la République (Paris: Assemblée Nationale, Commission des Lois, October 2018). Available at : *http://www.assemblee-nationale.fr/15/rap-info/i1335.asp*
14. Camille Gosselin, "La police prédictive : enjeux soulevés par l'usage des algorithmes prédictifs en matière de sécurité publique" (Paris: IAU Île-de-France, 2019). Available at: *https://www.institutparisregion.fr/fileadmin/NewEtudes/Etude_1797/Etude_Police_Predictive_V5.pdf*
15. Christophe-Cécil Garnier, "Dans tous les commissariats de France, on utilise la reconnaissance faciale," StreetPress, April 7, 2021. Available at: *https://www.streetpress.com/sujet/1617723420-tous-commissariats-france-utilise-reconnaissance-faciale-police-gendarmerie-justice-surveillance-zad-squat-libertes-societe*
16. URL to the clip: *https://video.lqdn.fr/w/xr4qR4AeBCa3JT3y7EWAjU*
17. Garnier, "Dans tous les commissariats de France, on utilise la reconnaissance faciale."

As well as the TAJ database, most immigration files now include photographs of faces, as does the *Titres électroniques sécurisés* database (TES, or Secure Electronic Documents). This was created in 2016 by the Ministry of the Interior and holds information about all applicants for identity cards and passports.[18]

Eventually, the Ministry of the Interior will be able to access the facial image of every person present on French territory, especially if it were to interlink its systems with similar databases in other European Union (EU) countries, and those operated by the EU itself.[19]

### 1.3.3 Illegal data collection and illegal database access

It is important to mention that illegal data collection is a common practice. Municipal police officers in a the Provence-Alpes-Côte d'Azur prefecture have used their personal smartphones to exchange sensitive personal data on private WhatsApp channels. This included video surveillance images, photos of people who had been stopped by the police, licence plates, identity documents and so on. While these practices are completely illegal, they are presumably commonplace, not just within municipal police forces, but also among the *Police Nationale*.[20]

In some cases, these practices are even partially institutionalised. In Cannes, for example, the *Police Municipale* have set up WhatsApp channels to enable local shopkeepers to report people suspected of theft or burglary.

Similarly, many municipal police forces are equipped with the Smart Police software sold by the start-up Edicia, which includes a 'predictive' module (discussed further in section 2). This software platform, used by hundreds of municipalities across the country, lets police officers use their phones or tablets to write up their reports 'on the move,' add photos, describe events or draw up reports. It also allows senior officers to monitor field teams from their offices, map incidents and receive various indicators in real-time.

18. La Quadrature du Net, "La reconnaissance faciale des manifestant·es est déjà autorisée", November 18, 2019. Available at : *https://www.laquadrature.net/2019/11/18/la-reconnaissance-faciale-des-manifestants-est-deja-autorisee/*
19. 'EU agencies and interoperable databases', *Statewatch*, https://www.statewatch.org/eu-agencies-and-interoperable-databases/
20. Éric Galliano, "Saint-Laurent-du-Var : Les policiers municipaux ont constitué leurs propres fichiers de délinquants," Nice Matin, November 20, 2023. Available at: *https://www.nicematin.com/justice/a-saint-laurent-du-var-les-policiers-municipaux-ont-constitue-leurs-propres-fichiers-de-delinquants-886441*

We obtained the user manual for the Smart Police system through a CADA request, have made it available on the Technopolice website.[21] One of the system's most important components its 'Field Activities' menu, which most *Police Municipale* officers use on a daily basis (see figure 1). This tool allows officers to register complaints, and to write and reference police reports documenting various offences that *Police Municipale* officers are authorised to report.[22]

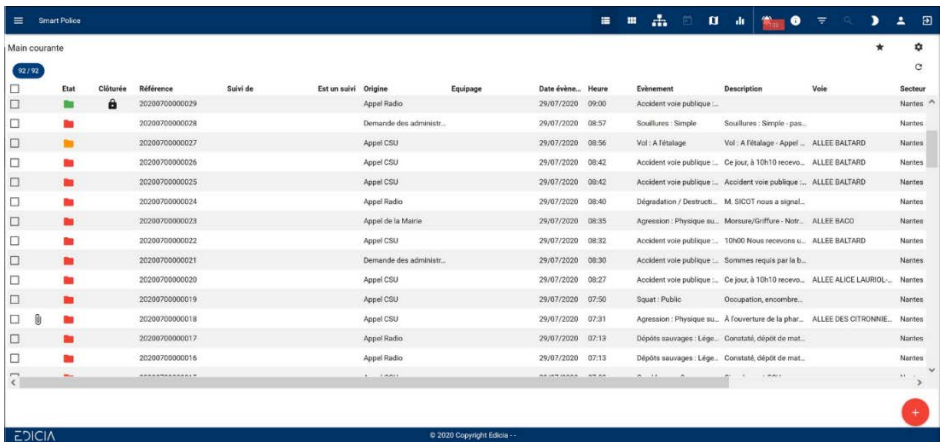



*Figure 1: A view of the 'Field Activities' module displaying the list of the latest complaints (with their origin, timestamp, etc.) (Smart Police User Manual).*

When registering complaints or incident reports, officers are required to provide, amongst other things:

- general information;
- the geographical location of the offence;
- the type of offence; and
- the identity and contact details of the suspect or witnesses (which can be recorded by scanning an ID card).

Within this application, there is another module that presents a major risk of illegal data collection: the *Demande administré* (citizen request) module. As its name suggests, this records reports made by citizens to the *Police Municipale* (including noise, damage, presence of a dangerous animal and more). This module can also be used to add geolocation data and photographs.

21. The Smart Police user manual is available at: *https://technopolice.fr/police-predictive/manuel-edicia/Edicia.html*
22. Officers are prompted to choose from a list of pre-selected offences drawn from a national database listing all types of offence (the NATINF database).

Finally, Smart Police includes a *Vigilance active* (active vigilance) module (see figure 2). Through this, users can upload unofficial information on past and future events (called 'alerts'). For example, if an important football game is scheduled for the following Sunday, a user can record this event. Similarly, if a police officer has heard a rumour about an unauthorised gathering (for example, on the street or via social media), they can create a new entry to record it. These various 'alerts' can then be transformed into 'missions' assigned to officers. This leads to the creation of new 'complaints' reports, but also feeds the *Analyse predictive* (predictive analysis) module of the Smart Police application, discussed in more detail later in this report.

These features of the Smart Police application present significant concerns regarding the recording of sensitive personal data in connection with unsubstantiated information. This would be illegal data collection. This feature also does not appear to comply with the regulatory framework for automated processing by municipal police forces managing police records, which prohibits the use of photographs.[23]

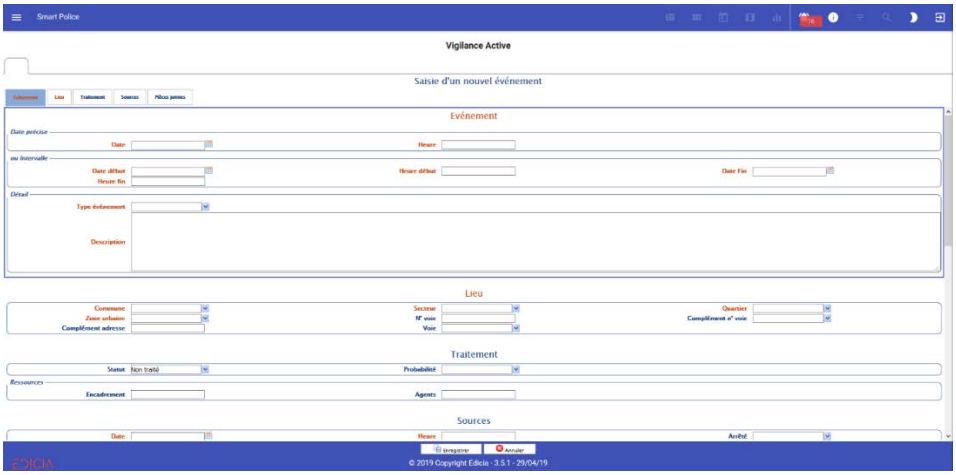



*Figure 2: Recording a new 'event' in the 'Active Vigilance' module (Smart Police user manual).*

23. "Arrêté du 14 avril 2009 autorisant la mise en œuvre de traitements automatisés dans les communes ayant pour objet la recherche et la constatation des infractions pénales par leurs fonctionnaires et agents habilités," accessed December 9, 2023, *https://www.legifrance.gouv.fr/loda/id/JORFTEXT000020692173*

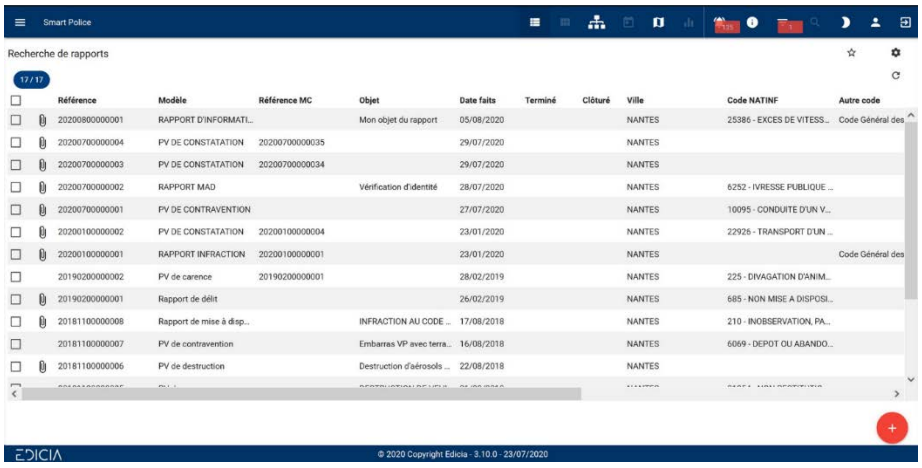



*Figure 3: A view of another feature available in the 'Activities" section displaying a list of all reports included in Smart Police (including complaints, reports, citizen requests, etc.) (Smart Police user manual).*

The *Inspection générale de la Police Nationale* (IGPN, or General Inspectorate of the National Police) has noted a significant increase in the number of cases of data misappropriation. There were 56 investigations for data misappropriation in 2022, compared with 38 in 2021 and 27 in 2020. The IGPN describes this trend as "worrying":

> The seriousness of these cases varies widely, depending on whether they are the result of "unhealthy" curiosity (e.g. accessing the files of an ex-girlfriend or a new partner, family members, a known person or a department head, without any money changing hands or any form of compensation) or of the trafficking of the information gathered for profit. These are the most sensitive cases, where confidential information from police databases is passed on to third parties, with (or without) a view to profit. While proof of illegal consultation is fairly straightforward for investigators, the same cannot be said for proof of compensation.[24]

The IGPN report states that "this situation is due both to the multiplication in the number of police files and to improved accessibility." This is due in particular to the growing deployment of NEO tablets and smartphones, which are specific mobile terminals used by the *Police Nationale* and *Gendarmerie*.[25]

24. Hippolyte Radisson, "L'IGPN relève une hausse de la "pratique préoccupante" du détournement de fichiers par les policiers en 2022," AEF.info, September 25, 2023. Available at: *https://www.aefinfo.fr/depeche/699730-l-igpn-releve-une-hausse-de-la-pratique-preoccupante-du-detournement-de-fichiers-par-les-policiers-en-2022*
25. Crosscall, "Crosscall equips the French Gendarmerie and Police with 200,000 terminals", March 15, 2021. Available at: *https://www.crosscall.com/en_FR/blog/neo.html*


*Figure 4: Administrative investigations into breaches of professional secrecy (light grey), including data misuse (blue) (excerpt from IGPN 2022 report).*

The IGPN believes that Artificial Intelligence (AI) could make it easier to detect illegal database access. However, we believe that the answer lies in a de-escalation and reduction of technologically-driven surveillance and security measures and infrastructure.

This means stopping the exponential increase in the population's data collection and the constant decline in the concrete guarantees to fundamental rights. The General Data Protection Regulation (GDPR) and associated legislation have, in many respects, contributed to this decline. This will be more effective than  obscure forms of technological solutionism. Regulatory bodies such as the CNIL should be doing their stated job, namely fighting abuses, rather than establishing *de facto* impunity for illegal data collection and access through their blameworthy *laissez-faire* attitude.

From this point of view, a first step in the right direction would be to ensure a detailed access record of *Police Municipale* users of Edicia. Sanctions should be imposed on police officers guilty of abuse. In a second phase, this type of system should be subject to strict regulation and systematic checks by the relevant authorities.

# 2. Overview of 'predictive' policing systems

In this section, we present findings on some of the 'predictive' policing systems used in France in recent years and highlight their common origins in criminological mapping.

## 2.1 The origins of criminological mapping

The French police can rightly claim to be pioneers in the field of so-called 'predictive' policing. As early as 1829, a young legal expert from the Ministry of Justice, André-Michel Guerry, set out to compile crime data for Paris. His *Compte général de l'administration de la justice criminelle en France* (General account of the administration of criminal justice in France)[26] was part of the world's first national system for centralising crime data.

This statistical database was founded on data collected quarterly in every French prefecture. It contained details of every criminal act brought before French courts, including age, sex, occupation of the accused, nature of the crime, and so on.

Guerry then turned his attention to criminological data and to the 'social attitudes' that he believed determined crime. He would later give birth to what is known as 'moral statistics', the discipline behind the development of criminology. Today, along with the sociologist Adolphe Quetelet, Guerry is considered criminology's founder.

In 1829, in collaboration with Venetian geographer Adriano Balbi, Guerry published a map showing the relationship between the level of education of the French population, and personal and property crime in France (see figure 5).

26. Michael Friendly, "The life and works of André-Michel Guerry, revisited," Sociological Spectrum 42, No. 4 6 (November 2, 2022): 233 59.

Today, this map is regarded worldwide as the first work of what would later be known in criminology as crime mapping. This is the starting point for 'predictive' criminology, with a colour code from light to dark indicating differences in crime rate. In this particular instance, his results showed that there was no statistical correlation between education and crime.

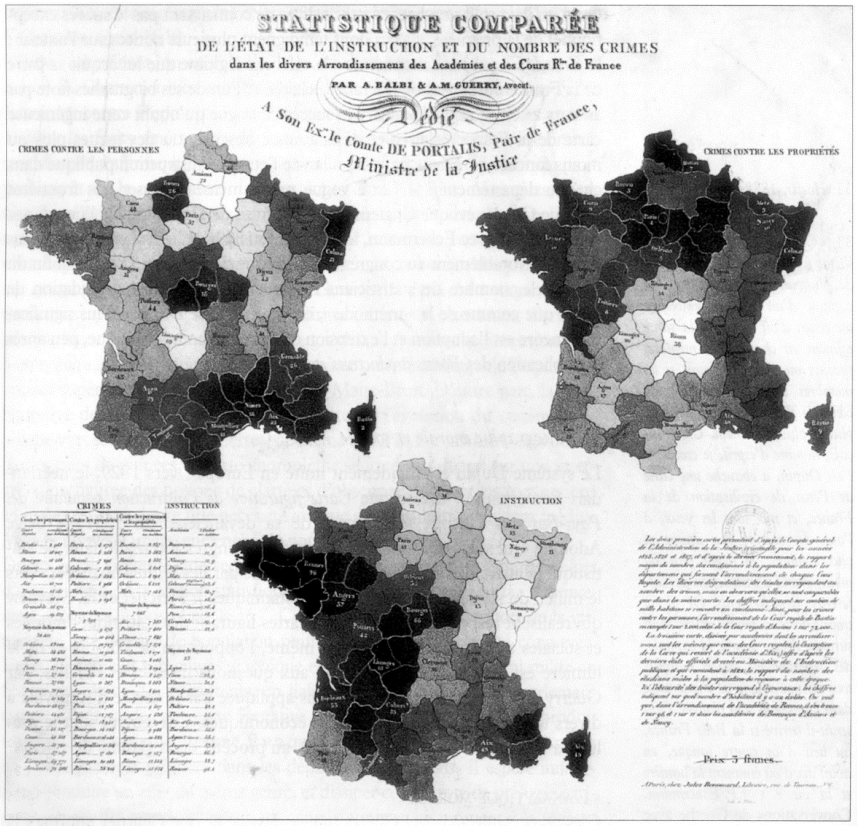



*Figure 5: Adriano Balbi and André-Michel Guerry (1829).* "Statistique comparée de l'état de l'instruction et du nombre des crimes dans les divers arrondissements des académies et des cours royales de France"*, Jules Renouard, Paris.*

In 1830, revolutionary events toppled the Restoration[27] and established the more liberal regime of the July Monarchy. The new regime appointed Guerry as Director of Criminal Statistics, a new department of the Ministry of Justice. In 1832, at the age of 29, Guerry published his *Essai sur la statistique morale de la France*, which is considered his major work. This essay contains many other criminological maps showing the relationships between crime and social and moral parameters: crime rates and wealth, suicide rates, donations to the poor, illegitimate births, age distribution of criminals, and so on. This work led to his growing recognition throughout Europe.

27. The period from 1815-1830 in France is known as the Bourbon Restoration. After Napoleon Bonaparte fell from power, the Bourbon family once again became monarchs of the country.

Guerry was also the inventor of the *ordonnateur statistique*, a machine that allowed him to establish relationships between statistical parameters.[28] The *ordonnateur* was based on the classical statistical methods of correlation[29] and regression analysis.[30] These would later be widely used in criminology, but were still underdeveloped at the time. The statistical comparisons made by Guerry using this method led him to establish relationships between types of crime, and various potential causes or associations.[31]

Since Guerry, criminology in France and around the world has continued to make extensive use of geographical approaches to crime.[32] This is particularly so in 'predictive' criminology. The systems we were able to analyse for this report are also based on geographical crime maps, with heat maps or 'hotspots' representing crime rates. Most of these systems also include the equivalent of Guerry's 'social and moral' parameters. These are now most often referred to as socio-economic or socio-demographic statistical indicators, such as:

- unemployment levels;
- school enrolment;
- level of education;
- number of nearby shops;
- average age;
- average household income;
- sex;
- nationality and immigration data;
- household composition.

However, whereas Guerry's research highlighted correlations on a regional scale, 'predictive' policing systems now focus on much smaller geographical units. As for the quantity of statistical data processed, and the speed at which these processes influence police practices, they are incomparable from Guerry's 19th century practices. This vastly increased quantity and speed can create and exacerbate feedback loops (self-reinforcing effects), with the potential to reinforce structural discrimination.

Finally, Guerry's results contradicted the prevailing cultural beliefs of the time – that crime was largely committed by uneducated and low-income people. This was determined largely by the moral panic of the bourgeois classes. By contrast, contemporary 'predictive' systems reinforce prevailing criminological postulates in their models, even though they have been largely invalidated by sociology.

28. Michael Friendly and Nicolas de Sainte Agathe, "André-Michel Guerry's Ordonnateur Statistique: The First Statistical Calculator?", The American Statistician 66, no. 3 (August 1, 2012): 195 200. Available at: *https://doi.org/10.1080/00031305.2012.714716*
29. See the "correlation" entry on Wikipedia: *https://en.wikipedia.org/wiki/Correlation*
30. See the "regression" entry on Wikipedia: *https://en.wikipedia.org/wiki/Regression_analysis*
31. See also the Larousse encyclopedia entry for Guerry: *https://www.larousse.fr/encyclopedie/personnage/Andr%C3%A9_Michel_Guerry/179996*
32. Melina Germes, "Cartographies policières : la dimension vernaculaire du contrôle territorial," EchoGéo, no. 28, July 8, 2014. Available at: *https://journals.openedition.org/echogeo/13856?lang=fr*

## 2.2 Using data and statistics to direct police action

Other statistical approaches have also been developed with the aim of optimising the police management of urban public space. From the 1990s onwards, as the idea of "new public management" gained ground in the police establishment, the criminological tools experimented with by United States police forces were gradually integrated into the practices of the French police. The New York Police Department (NYPD) under the leadership of William Bratton is particularly relevant. COMPSTAT software initially used by the NYPD was deployed by police commanders in Paris as a guiding tool.[33]

Initially designed to collect crime statistics in order to improve police responsiveness and anticipation, COMPSTAT rapidly became a key instrument in a policy of setting target numbers. This was part of an attempt to legitimise the actions of police bureaucracies in the administrative and political spheres. Sociological studies have shown how the quantification of police activity has made it possible to place responsibility on every agent in the chain of command, from the police commander to the officer on patrol. According to Emmanuel Didier:

> This adoption of neoliberal management methods and objectives is radically transforming this core government activity. [...] What is valued by the top of the hierarchy is precisely initiatives taken by officers, all officers, a kind of reformatted autonomy; initiatives which are not constrained by rules but encouraged by a sense of responsibility, a responsibility which is, incidentally, very different from the classic administrative responsibility of law enforcement officers.[34]

The consequences of this policy have been widely documented by police sociologists: officers at the lower levels of the hierarchy attempted to reduce the number of cases in order to avoid pressure from the 'top'. This resulted in a refusal to collect complaints, sending complainants from one police station to another, or regrouping or re-characterising recorded facts.

It led, above all, to a concentration of police activity on the most 'profitable' offences. The aim was to increase the number of 'solved' cases, to the detriment of more complex cases. This leads to differentiated treatment of 'lower-class' crime – linked, for example, to the possession and consumption of narcotics – compared to white-collar crime.[35]These rationales mirror the trends underlying contemporary 'predictive' policing technologies.

33. Emmanuel Didier, "'Compstat' à Paris : initiative et mise en responsabilité policière", Champ pénal/Penal field, no. Vol. VIII, June 11, 2011.
34. Ibid.
35. Laurent Bonelli, "Les modernisations contradictoires de la police nationale", in L'État démantelé, Cahiers libres (Paris : La Découverte, 2010), 102, 17. Available at: *https://doi.org/10.3917/dec.bonel.2010.01.0102*

## 2.3 'Predictive' policing systems

In recent years, criminological mapping and the use of statistical indicators to guide police action took the form of location-focused crime 'prediction' systems. These now seem to have been abandoned, or are at least of minor importance. This can be seen with the MapRevelation system, discussed later. This section considers six systems in use or formerly used by police in France.

### 2.3.1 Risk Terrain Modelling (RTM) (Paris Police Prefecture)

Risk Terrain Modeling (RTM) is a methodology developed by two academics at Rutgers University in the USA. A geo-statistician at the former *Observatoire National de la Délinquance et des Réponses Pénales* (ONDRP, or National Observatory of Criminality and Penal Responses)[36] worked for several years the university to test this methodology and the crime 'prediction' algorithm it produced. The algorithm is used by the *Direction de la Sécurité de Proximité de l'Agglomération Parisienne* (DSPAP, or the Paris Region Local Security Division), which is part of the Paris Police Prefecture (National Police) and is based on DSPAP crime data.

By relying on the correlation between criminality and 'environmental variables' rather than socio-demographic variables, RTM perpetuates the logic of situational crime prevention.[37] A police officer told the academic Camille Gosselin:

> I have been using the RTM (Risk Terrain Modelling) algorithm for over eight years [meaning they starting using it in 2011 or 2012]. [...] It is now a web application, much more powerful than it previously was, but it uses the same principle: it's in line with situational crime prevention, meaning that we identify the contextual and environmental elements that make crime happen. The causes of crime are inequalities, poverty, unemployment, genetics, etc., as well as a favourable environment in the medical sense of the term. So by analysing an environment and identifying the factors that aggravate risk, we can predict what might happen in a similar environment ...[38]

36. Following its closure in 2020, part of the Observatoire's activities have been transferred to the French Ministry of the Interior's Service Statistique Ministériel de la Sécurité Intérieure (SSMSI, part of the Institut des hautes études du ministère de l'Intérieur or IHEMI, created in September 2020).
37. Situational crime prevention is a set of methods designed to take safety into account when planning public or private spaces. Its aim is to reduce people's sense of insecurity. The concept was first coined by urban planning professionals and researchers in the English-speaking world in the name of 'crime prevention through environmental design', before being imported to Europe (source: Wikipedia).
38. Camille Gosselin, "La police prédictive : enjeux soulevés par l'usage des algorithmes prédictifs

To do this, the RTM system used by the Paris police draws on data from the *Logiciel de Rédaction des Procédures de la Police Nationale* (LRPPN, or *Police Nationale* Procedure Writing Software). This compiles localised data on crime occurrences that are then linked to common environmental variables for each type of crime. By deducing supposed correlations between environmental factors and the occurrence of criminal acts, the RTM model aims to anticipate crimes, and prevent them by allocating appropriate resources.

The algorithm assigns a vulnerability value to locations, based on the types of crime selected and 'situational' characteristics, including:

- the amount of street lighting;
- metro or bus stations;
- outdoor seating areas of cafés;
- fast-food outlets;
- public toilets;
- pharmacies;
- grocery stores;
- bars;
- trees and benches;
- certain types of shops;
- schools;
- post offices.

Despite the allusions in the quote above, it does not appear that the RTM tool includes socio-demographic data for selected geographical areas (see figures 6 and 7).[39] However, the data it uses on an area can act as a proxy for socio-demographics, such as the presence and density of fast-food restaurants, or public housing.

In 2018, Jean-Luc Besson, head of the ONDRP's geostatistical department, claimed that RTM had a real impact on crime rates. He explained the logic of RTM systems:

> The logic is that if there is a concentration of incidents in a particular place, it's because there's a favourable environment, and to look for the reasons for this predisposition … If we look at pickpocketing near cash machines, we'll ask ourselves: are the most affected cash machines open day and night? Are they near a crossroads or a train station, etc.?[40]

---

en matière de sécurité publique" (Paris: IAU Île-de-France, 2019). Available at: *https://www.institutparisregion.fr/fileadmin/NewEtudes/Etude_1797/Etude_Police_Predictive_V5.pdf*

39. Alejandro Giménez-Santana, Leslie W. Kennedy, and Joel M. Caplan, "Risk terrain modelling and the study of the physical determinants of criminal behaviour", Cahiers de la sécurité et de la justice, no. 47 (March 2019), 126 34. Available at: *https://rscj.newark.rutgers.edu/research/publication/gimenez-santana-a-kennedy-l-w-caplan-j-m-2020-risk-terrain-modeling-and-the-study-of-the-physical-determinants-of-criminal-behavior-cahiers-de-la-securite-et-de-la-justice-revue-de-li/*
40. Thibault Sardier, "Cartographie criminelle : surveiller et prédire," lemonde.fr, January 5, 2018.

| Environmental Factors | N | Operationalization | Spatial Influence |
|---|---|---|---|
| Grocery stores | 1260 | Proximity and Density | Up to 450 meters (or up to 3 increments of 150 meters) |
| Bars and Nightlife | 4448 | Proximity and Density | |
| Low-cost Cafes (Café à 1 Euro) | 178 | Proximity and Density | |
| Car-sharing stations | 113 | Proximity | |
| Shops and Malls | 36 | Proximity | |
| Coffee Shops | 3443 | Proximity and Density | |
| Banks | 2712 | Proximity and Density | |
| Terrace / Outside seating areas / Stands | 20893 | Proximity and Density | |
| Fast food restaurants | 113 | Proximity and Density | |
| Gas Stations | 144 | Proximity and Density | |
| Post Office | 210 | Proximity | |
| Public Housing | 2891 | Proximity and Density | |
| Movie Theaters | 86 | Proximity and Density | |
| Parks and Public Gardens | 511 | Proximity and Density | |
| Pharmacies | 1037 | Proximity and Density | |
| Coin-operated public toilets (Sanisette)<br>Schools<br>Metro Stations<br>Train Stations (RER and SNCF)<br>Taxi Stations<br>Tourist Areas[1]<br>Museum<br>Bike-sharing Stations (Vélib) | 395<br>1593<br>246<br>72<br>120<br>12<br>37<br>68 | Proximity<br>Proximity and Density<br>Proximity and Density<br>Proximity and Density<br>Proximity<br>Proximity and Density<br>Proximity and Density<br>Proximity and Density | |

*Figure 6: List of potential risk factors according to the RTM algorithm.*

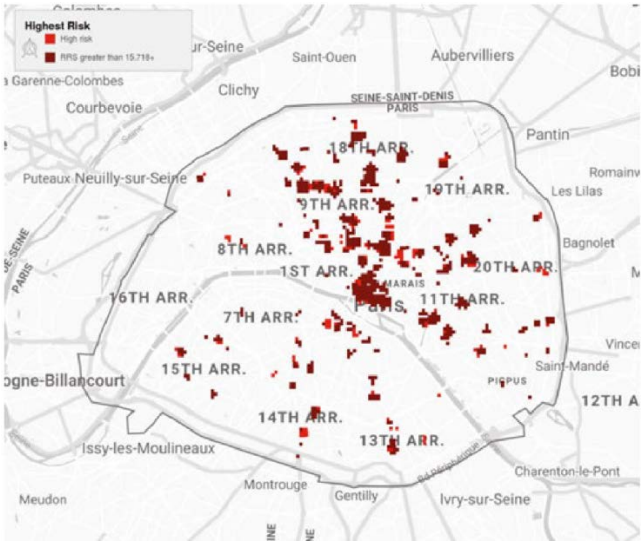



*Figure 7: The areas marked red are considered 'at risk'. These appear to be areas with high traffic flow or tourist locations.*

However, the input data is not always easy to use, according to a source with knowledge of the system:

> It's administrative data, but it's not used operationally or analytically. Analytical data requires a certain number of variables and information. [...] Even the location is not always given in the data, sometimes it doesn't appear at all, nor does the address, or there's just a mention of 'Paris.' But Paris isn't a place, it's too big! The notion of place, the geographical location of the intervention, just that alone is complicated![41]

Available at: *https://www.lemonde.fr/idees/article/2018/01/05/cartographie-criminelle-sur-veiller-et-predire_5237723_3232.html*

41. Camille Gosselin, "La police prédictive : enjeux soulevés par l'usage des algorithmes prédictifs

The subsequent analysis of the system's results by police officers also seems inadequate:

> We [the police] are mostly concerned with reporting, we ask them [the algorithms] for statistics, they spit them out, but they do very little analysis of this data to understand the phenomena. We see that it's going up, that it's going down, where it's going up, where it's going down, but there's no analytical structure behind it, and we never answer why.

In this source's opinion, the potential effects of sending patrols to a given area on the geographical displacement of crime are not really evaluated. The source indicated that, like the other systems studied in this report, the RTM methodology allows institutions using it to evade any reflection on the structural causes of crime – an aspect to which we will return in the analytical section.

### 2.3.2 MapRevelation (Sûreté Globale, Police Municipale)

In 2020, we became aware of one of the first 'predictive' policing systems deployed in France: MapRevelation,[42] marketed by the company Sûreté Globale.[43] The company had provided the system to municipal police forces in Montpellier, Lyon, Lille, Villeurbanne, Montauban, Angers, Colombes and Melun Val de Seine.

At the time, we were able to obtain documentation on MapRevelation from the city of Montpellier.[44] The company's website also indicated other uses. The border police had used the software in 2010 to "produce statistical dashboards on the flow of illegal migrants."[45] The Paris Police Prefecture had used it in 2011 to "determine areas of criminal concentration" on New Year's Eve. In 2015, the *Gendarmerie Nationale* deployed the system within its judicial centre.[46] This was for "criminal analysis of complex cases, and as part of its decision-making support for senior staff, particularly in terms

en matière de sécurité publique" (Paris: IAU Île-de-France, 2019). Available at: *https://www.institutparisregion.fr/fileadmin/NewEtudes/Etude_1797/Etude_Police_Predictive_V5.pdf*

42. MapRevelation website: *https://web.archive.org/web/20230223042353/https://www.maprevelation.fr/*
43. Sûreté Globale website: *https://web.archive.org/web/20230610153444/https://www.sureteglobale.org*
44. La Quadrature du Net — Technopolice, "La police prédictive progresse en France. Exigeons son interdiction !"
45. See *https://web.archive.org/web/20230203040444/https://www.sureteglobale.org/2010/11/la-direction-centrale-de-la-police-aux-frontieres-choisit-map-revelation/*
46. The Pôle Judiciaire de la Gendarmerie Nationale (PJGN, or the Judicial Centre of the Gendarmerie Nationale).

of anticipation and forecasting." The company's website also displayed the many logos of its partners. These were either MapRevelation users or "research and development" partners who had contributed their data or expertise to the training, improvement and development of the software.[47]

Since 2023, we have been taking a renewed interest in MapRevelation. Sources have indicated to us that the system was designed by Christophe Courtois, manager of the company Sûreté Globale.[48] MapRevelation is also said to have been trained on a 'terrorism' database to predict attacks,[49] as well as data held by the authorities using it (*Police Municipale*).

Like Risk Terrain Modelling, MapRevelation's algorithm appears to be based on the premise of environmental determinism, whereby environmental factors in an area are thought to be a major contributing factor of crime. It is also based on the theory of situational crime prevention, which takes various factors into account. These include the presence of so-called 'motivated' individuals, a 'criminogenic' environment, and the absence of police and other potential crime 'deterrents'. The other criminological theory underpinning the model is the regularity of the crime phenomenon. This is inferred using geographic data of past crimes and a statistical model fed by training data.

The system was designed to be used as a strategic guidance tool for police forces, linked to 'hypervisors', software that allows for multiple virtual machines to run on a single physical machine, located in security force command centres. MapRevelation's output data consists of geographical maps with geolocated crime prediction and heat maps (see figure 8).

47. See the list of Sûreté Globale customers and partners: *https://web.archive.org/web/20230610143211/http://www.sureteglobale.org/works/*
48. See *http://www.sureteglobale.org/en/equipe.htm*
49. International terrorism database: *https://www.start.umd.edu/gtd/access/*. It shows that France only has 20 entries from 2016 to 2019, which leaves one wondering about the relevance of this dataset.

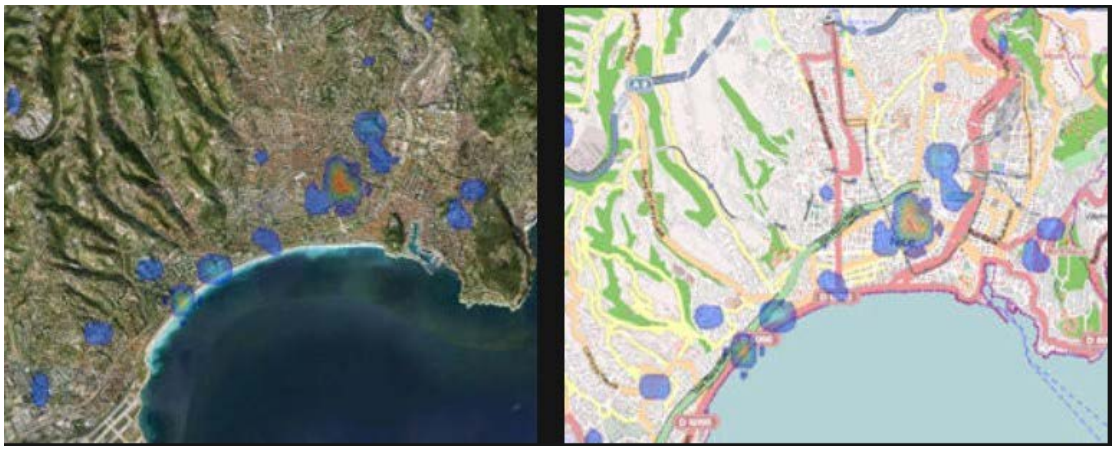


*Figure 8: A screenshot of MapRevelation (source: Sûreté Globale website).*

The maps produced by the model display 'heat zone' predictions for the next 24 hours, with blue-ringed zones indicating the locations where there is supposedly the greatest likelihood of crime occurring.

Despite Sûreté Globale's many customers and business partners, the company has experienced major setbacks. In 2018, it was placed under receivership for cessation of payments, with a 10-year restructuring plan. The company's website and social media accounts, although still in existence, no longer seem very active. Apart from its founder-manager Courtois, its seems that there is no one else working at the company. We do not know the reasons that led to the company's insolvency, nor do we know at this moment in time whether its contracts continued to be honoured. In May 2024, the company closed down.[50]

Our most recent requests for access to documents from some of the local authorities that have used MapRevelation have not been answered.[51]

## 2.3.3 Predvol (Etalab, Gendarmerie Nationale)

PredVol is a 'predictive' policing system developed from 2015 onwards mainly by Florian Gauthier, a young data scientist temporarily hired by the French government's Etalab public data programme. Etalab sits within the *Direction interministérielle du Numérique* (DINUM, or the Inter-ministerial Digital Directorate). The name PredVol seems to allude to the US predictive policing system and company Predpol.[52]

50. Pappers Enterprises: *https://www.pappers.fr/entreprise/sgo-sureteglobaleorg-511724866*
51. See the CADA requests sent to local authorities about MapRevelation: https://madada.fr/search/Map%20Revelation/all.
52. PredPol was renamed Geolitica, and then partially acquired by another predictive policing

The *Gendarmerie* and the *Police Nationale* were not involved in PredVol's development. Literature on PredVol, including its algorithms, operational use and test results is sparse.[53] Accorting to Florian Gauthier, PredVol was tested for six months in 2016 by the *Police Nationale* and the *Gendarmerie*, in the Oise prefecture. The algorithm was only designed to predict car thefts, an offence for which the reporting rate to law enforcement authorities is generally high (over 90% of cases are reported to the police).[54]

During the system's pilot, PredVol offered senior police officers a daily forecast of the risk of car theft, a map of car theft history and a "typology of neighbourhoods according to the nature of the offences committed there," according to Gauthier. The main purpose of the software was to guide police patrols and help them to better deploy their resources in predicted risk areas. Gauthier justifies the need to develop car theft 'prediction' software through 2014 data from the Oise prefecture that demonstrates discrepancies between patrol zones and car theft zones (see figure 9). According to Gauthier:

> A quick examination reveals that some areas, which are heavily patrolled by law enforcement agencies, also see a large number of car thefts (Zone A), while others, although heavily affected by car thefts, are rarely patrolled (Zone B). To what extent would it be possible to anticipate car thefts in order to better direct *Police* and *Gendarmerie* patrols?[55]

company, SoundThinking. See: WIRED, "The Maker of ShotSpotter Is Buying the World's Most Infamous Predictive Policing Tech", September 27, 2023

53. The following evidence is based on these sources: Florian Gauthier, "Prédire les vols de voitures ?", Etalab (blog), January 12, 2018, available at: *https://www.etalab.gouv.fr/predire-les-vols-de-voitures;* Camille Gosselin, "La police prédictive"; "Sous le capot de la police prédictive," Courrier Picard (blog), April 2, 2018. Available at: *https://web.archive.org/web/20220301133732/https://www.courrier-picard.fr/art/88625/article/2018-02-04/sous-le-capot-de-la-police-predictive*
54. Yann Lecorps, Gaspard Tissandier, "PAVED with good intentions: an evaluation of the Gendarmerie predictive policing system", Centre d'Économie de la Sorbonne (CES), Université Paris 1 Panthéon Sorbonne (Paris, September 2022).
55. Gauthier, "Prédire les vols de voitures ?"

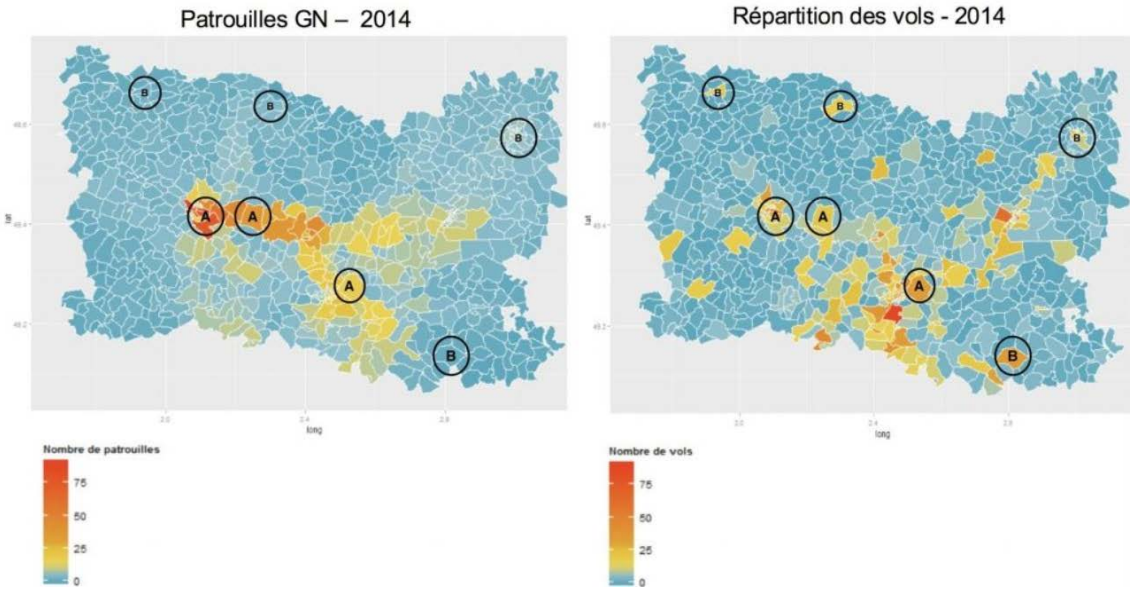



*Figure 9: Maps provided by Gauthier showing discrepancies between the distribution of police patrols (left) and car theft zones (right) (Source:* Prédire les vols de voitures?*).*

**Input data**

There were two main sources of input data for PredVol.

Firstly, the complaints databases of the LRPPN (*Logiciel de Rédaction de Procédures de la Police Nationale*) and LRPGN (*Logiciel de Rédaction de Procédures de la Gendarmerie Nationale*). These are the IT systems used to record all complaints and reports from the police's judicial and administrative proceedings. They record information on all crime events, as well as information surrounding the events, that may be required for judicial processing (or other forms of processing).[56]

PredVol extracts data from these databases about car theft events in the last three years. Each car theft event is recorded in a row on the database with the following information: XY coordinates where the theft occurred, date, and sometimes additional information on the stolen vehicles such as make, model, colour, and so on.

56. Decree no. 2011-110 of January 27, 2011, authorising the creation of an automated personal data processing system called Logiciel de Rédaction des Procédures de la Police Nationale (LRPPN), January 2011: *https://www.legifrance.gouv.fr/loda/id/JORFTEXT000023491308/*

Second, the French geographical cartographic division system IRIS, which is maintained by the *Institut National de la Statistique et des Études Économiques* (INSEE, or the National Institute of Statistics and Economic Studies).[57] According to the INSEE, IRIS divides each French municipality into neighbourhoods of around 2,000 inhabitants. This forms the basis for the geographical distribution of intra-municipal statistics.

In PredVol, information from the LRPPN and LRPGN databases is cross-referenced with the official IRIS map of administrative divisions. According to Gauthier, IRIS was used to enrich the training dataset with over 600 socio-demographic variables on these areas, including unemployment rate, school attendance, number of nearby shops, and average population age.

Finally, PredVol's input data included indicators on the time of day when thefts occurred. Gauthier gives the following examples:

> Was there a theft the day before? Or the day before that? In nearby neighbourhoods? What was the weather like that day? The idea is to bring as many variables as possible into our database, without any preconceived ideas, and then let the machine learning algorithms select the best predictors to anticipate car thefts.

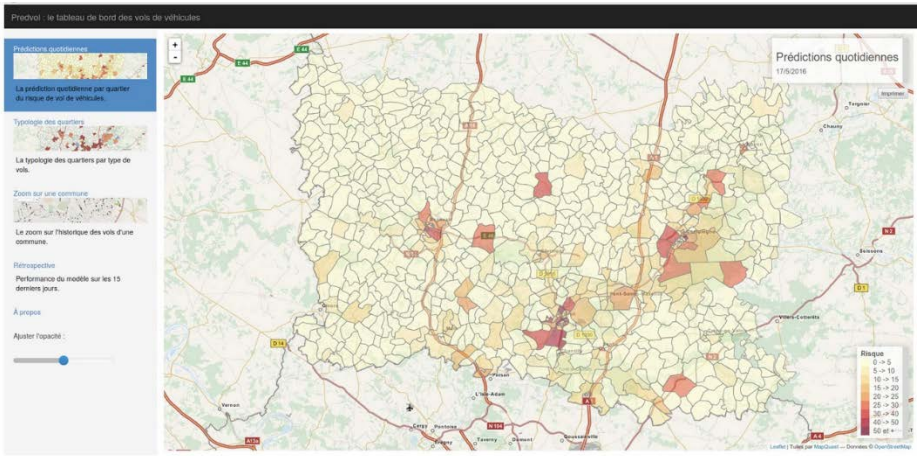
As we discuss later, the 600 socio-demographic indicators from IRIS were also used for the PAVED system. The developers of PredVol and PAVED explained that these are open-data statistics that are regularly updated by the INSEE. As we explain in the analytical section, the presence of these socio-demographic variables increases the likelihood of groups already impacted by structural discrimination being further targeted by police. We were unable to find complete datasets covering all these indicators,[58] raising questions about the quality of the input data, and in particular their chronological and spatial accuracy.

57. INSEE, Découpage infra-communal IRIS, available at: *https://www.INSEE.fr/fr/information/2017499*

58. Looking at the open data provided by the INSEE on its website, we were unable to find a single dataset containing all 600 indicators. We were able to access via the website the French unemployment rate dataset which is updated every three months. On another INSEE webpage, we found a list of indicators displayed interactively on a map of France but this did not amount to 600. We have made several CADA requests to both the INSEE and the Ministry of the Interior to obtain precise information on these 600 variables. The INSEE replied claiming that it knew nothing about their use in 'predictive' policing, and relayed our request to the Ministry of the Interior, who ignored it. We referred the matter to the CADA, who issued a favourable opinion, which the Ministry of the Interior once again ignored.

### Training, experimentation, discontinuation

According to Gauthier, the input data was fed into three different machine learning algorithms. One of which was apparently the same type of algorithm as that used by PredPol in the USA. In the end, only one of these test algorithms was selected for implementation in PredVol: the predictive heat map algorithm. This is simpler but supposedly produces equally valuable results as the more complex PredPol-type algorithms. Apart from the general descriptions provided by Gauthier, we have found no detailed public information on the mathematical machine learning model used by PredVol. No software code, evaluation datasets or evaluation results have been made public.




*PredVol – Prédictions quotidiennes*

*Figure 10: A screenshot of PredVol's daily forecasts (source:* Prédire les vols de voitures ?*).*

In 2016, the software was tested for six months by the *Police Nationale* and the *Gendarmerie* in the Oise prefecture, and in particular by the police force's *Brigade Anti Criminalité* (BAC, or anti-crime brigade) patrols. PredVol was designed for mobile use during police patrols, and was directly available on police officers' tablets in the test area.

The system was later abandoned. It appears that it merely confirmed the field experience of police officers and gendarmes, who had identified the supposed 'high-risk' areas themselves some time ago. In her report on predictive policing,[59] researcher Camille Gosselin cites several passages from interviews with officers who had used the software. They described it as a

59. Camille Gosselin, "La police prédictive : enjeux soulevés par l'usage des algorithmes prédictifs en matière de sécurité publique" (Paris: IAU Île-de-France, 2019). Available at: *https://www.institutparisregion.fr/fileadmin/NewEtudes/Etude_1797/Etude_Police_Predictive_V5.pdf*

"ridiculous experiment" and a "painful experience" that produced "meaningless and unscientific results". "[I]n the end" explained one source, "the region was divided into administrative divisions that had no geographical relevance whatsoever; it wasn't meaningful for the officers, and it wasn't accurate either."

In his article, Gauthier concludes that the 'predictive' analysis feature was subsequently removed from the software. It became a simple tool showing statistical crime maps. The software, renamed mapVHL, was then supposedly integrated with other technologies used by police forces and is now thought to be widely used in the field. However, the only mention of mapVHL that we have found is the one in which Gauthier makes this conclusion. He has not worked for Etalab since 2018.

### 2.3.4 PAVED (Gendarmerie Nationale)

As early as 2015, while PredVol was being developed within Etalab, the *Gendarmerie*'s *Service Central de Renseignement Criminel* (SCRC) conducted research into AI algorithms and models for 'predictive' analysis. This led to the development of PAVED in 2017. PAVED is an acronym for *Plateforme d'Analyse et de Visualisation Évolutive de la Criminalité* (Evolving crime analysis and visualisation platform).

PAVED, a decision analysis platform, allegedly uses AI to 'predict' two types of crime – vehicle theft and burglary – at street level for the following few days, for a maximum of up to a week.

At the end of 2017, PAVED was deployed in 11 prefectures in the '*Gendarmerie* zone'. This test phase was meant to last until early 2019, before nationwide deployment. However, this plan did not go ahead: PAVED was never deployed nationwide. The project was put on hold pending a potential "reuse," according to Colonel Patrick Perrot. Perrot is PAVED's main designer, and now a coordinator for AI within the *Gendarmerie*, as well as AI advisor to the Ministry of the Interior's cyberspace commander.[60]

60. See: *https://www.linkedin.com/in/patrick-perrot/*

## The origins of the project

Colonel Perrot published the first research article on PAVED in 2014. In the article, he justifies the need for crime 'prediction' tools:

> To anticipate, predict and act remains one of the objectives of any structure confronted with an evolving phenomenon. Crime, inherent to human history, is constantly evolving, even mutating. In the field of crime analysis, the scientific approach enables us to develop modelling techniques capable of understanding and preparing for future developments. The notion of anticipation is now a determining factor in the field of crime.[61]

The research article is Perrot's attempt to mathematically model the social reality of crime. His approach sets out to detect crime by describing the precursor signs of an event. These weak signals (for example, the introduction of a new type of narcotic or an increase in the price of copper) are modelled using a mathematical approach.[62] Secondly, he attempts to understand the supposed causes of crime by analysing potential variables and measuring their influence on the evolution of the observed event. This is done using mathematical linear regression methods:

> [W]e can explain burglaries (endogenous variables) on the basis of socio-economic variables such as, for example, the age of the population, average household income, and level of schooling ('exogenous' variables).

Finally, he seeks to anticipate crime by forecasting the temporal evolution of particular types of offence through the mathematical study of time series. What is striking is that, despite these theoretical considerations, Perrot offers no convincing proof or explanation, merely statistical correlations. Nor does he refer to any previous research along these lines. In doing so, as we explain below, he runs the risk of confusing correlation with causation. He also relies on socio-demographic variables that may amplify structural discrimination.

In another scientific article published in 2017,[63] Perrot revealed the first

61. Patrick Perrot, "L'analyse du risque criminel : l'émergence d'une nouvelle approche", Revue de l'Électricité et de l'Électronique REE 2014-5 SEE (December 1, 2014). Available at: *https://www.researchgate.net/publication/274071556_L'analyse_du_risque_criminel_l'emergence_d'une_nouvelle_approche*. See also: Patrick Perrot, "Le renseignement criminel : de nouvelles perspectives contre la criminalité organisée", Revue Défense Nationale 779, no 4 (2015) : 11 15.
62. Signal theory, *Ibid.*
63. Patrick Perrot, "What About AI in Criminal Intelligence? From Predictive Policing to AI Perspectives," European Law Enforcement Research Bulletin, no. 16 (August 14, 2017): 65 75. See also: Patrick Perrot, Valescant, Nicolas, and Camara, Daniel, "Forecasting criminal patterns for decision making," URSI France Journées scientifiques 2017 (February 1, 2017). Available at: https://

screenshots of the software. His explanations become clearer as he gives more details on the method adopted by PAVED. He explains that the model uses various spatial statistics and time series methods: exponential smoothing, autoregressive moving average and neural networks (also known as machine learning). Perrot also gives information on the two databases used to train and evaluate the mode: crime data from 2008-2014 to train the model, and data from 2014 to evaluate the model. The article proposed crime 'forecasts' for the year 2015.

**Tactical media publicity upon launch**

In 2017, the name PAVED had not yet appeared in scientific papers. In early 2014, the name PredPol, a software used by various police forces in the US was first mentioned in French media. 'Predictive' policing was the subject of criticism in documentaries and press articles.

In 2018, when the *Gendarmerie* began experimenting with PAVED 'in the field', they publicised the system widely in various media. They used terminology that distanced the system from the negatively charged term 'predictive policing'.[64] Instead, their publicity used more 'neutral' or 'technical' terms such as "criminal risk analysis", "predictive analytical method", "predictive analysis," "decision-support tool," or "statistical evaluation of future crime rates in a defined area." They highlighted how the system provided information that supported officers with daily planning decisions.

While this is a more accurate description of how the system functions, it is a long way from Perrot's research articles, which specifically discuss a predictive logic.

**Input data**

PAVED's input data includes burglaries and car thefts over the previous nine years, sourced from the ONDRP,[65] whose activities were transferred in 2020 to a department within the *Institut des Hautes Études du Ministère de l'Intérieur*. In addition, the system draws upon 687 socio-demographic indicators sourced from the INSEE, the same source as PredVol.

In a 2016 article, Perrot explains that only 15 out of the 687 indicators were selected to feed PAVED. According to Perrot, the INSEE data was from 2011–2012 and therefore not entirely up to date, hence the limited choice of just 15 variables.

www.ursi-france.org/fileadmin/journees_scient/docs_journees_2017/data/articles/000054.pdf.

64. See, for example, "22 v'là la police prédictive!", La méthode scientifique, France Culture, December 2018. Available at: https://www.radiofrance.fr/franceculture/podcasts/la-methode-scientifique/22-v-la-la-police-predictive-1641835.

65. The National Observatory of Criminality and Penal Responses

Again, we were unable to establish the exact list of these variables, but we do know that they include the following indicators, among others:

- population;
- age;
- gender;
- nationality and immigration data;
- household income;
- household composition; and
- level of education.

These variables were determined using mathematical regression methods. The use of such data can operate as a proxy for ethnicity or race. This poses a high risk of reinforcing the targeted policing of low-income and racialised people and the neighbourhoods in which they live.

**Interface**

PAVED consists of two windows, as shown in the figure below.

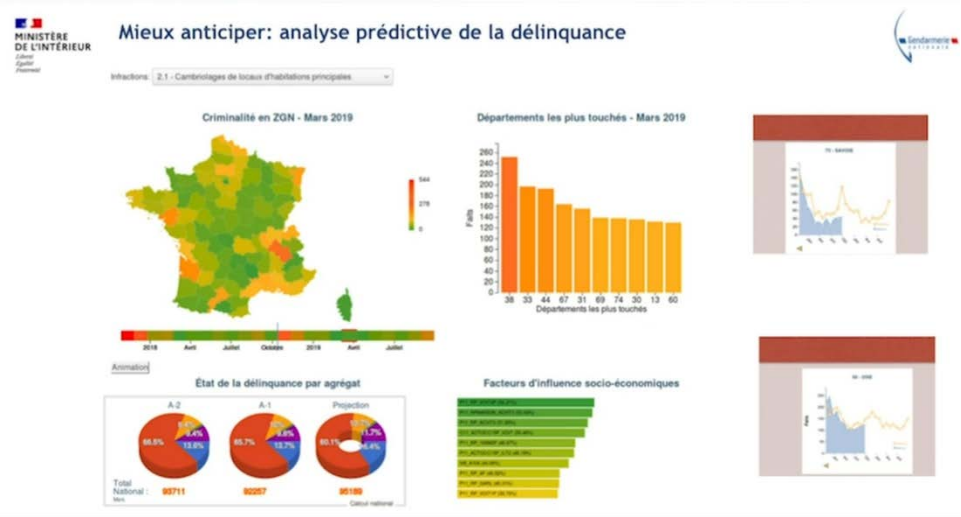



*Figure 11: Screenshot of the lecture given by Colonel Patrick Perrot, Colloque de l'Institut de Droit Privé—UT Capitole—September 8, 2021 (source: 'AI et enjeux de sécurité," at 23m10s).*

The macroscopic view seems primarily designed to support *Gendarmerie* with making decisions about how many officers to deploy in each zone. Within this image, the map on the top left shows a each crime type selected in the "offences" field, with a heat map showing risk 'predictions' for crime rates (red fwior high rates, green for low). The histogram in the top middle of the image displays the geographic divisions most likely to be affected by the crime type in question. The histogram in the bottom middle displays the socio-demographic variables considered most influential for the crime type.

The second view in PAVED shows daily predictions (see figure 12). This view is useful for determining which zones in a selected geographic area are supposedly the 'highest risk' for a selected period of time. According to Perrot, the 24-hour forecasts are not very reliable: good 24-hour predictions would require a daily refresh of PAVED's input data. This is impossible with the current state of technology and criminal proceedings. There is always a delay between an offence being committed, the data being registered by the police and its inclusion in the PAVED database. In Perrot's opinion, the system's best predictions are those obtained for the following week.

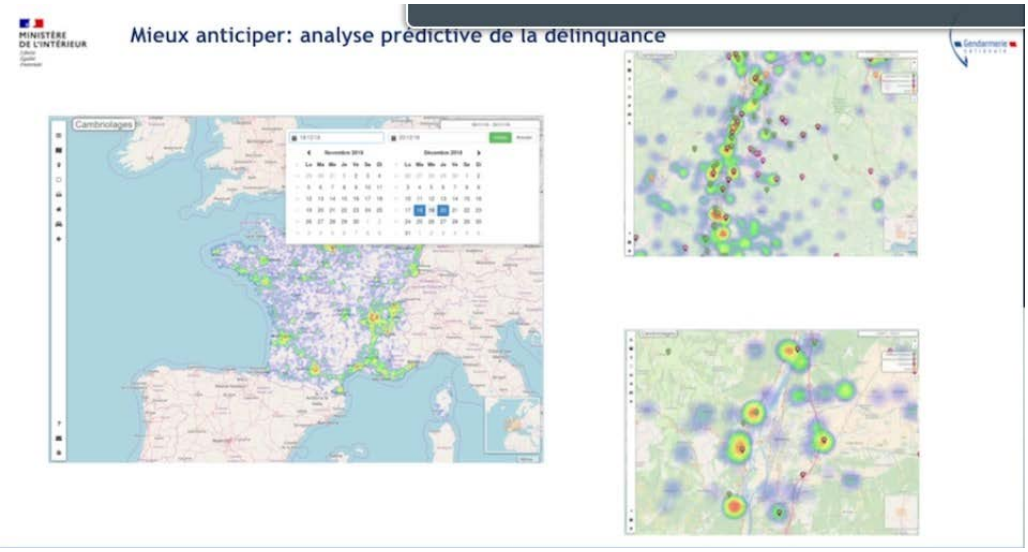



*Figure 12: Screenshot of the lecture given by Colonel Patrick Perrot, Colloque de l'Institut de Droit Privé—UT Capitole—September 8, 2021 (source: 'IA et enjeux de sécurité," at 26m10s).*

**Deployment**

PAVED was tested in 11 prefectures by the *Gendarmerie,* for 18 months (October 2017 to April 2019). Following the pilot, the software was due to be rolled out nationwide. However, this did not happen. Unlike PredVol, which was available on tablets for every officer, PAVED was used exclusively by *Gendarmerie* operational commanders. Based on the system's results, they were required to determine how many officers to send to each zone for the following days. We have no information from the Ministry of the Interior or the *Gendarmerie Nationale* with evaluation results from this trial period, nor any indication as to why the planned roll-out did not go ahead.

According to one source, PAVED was used to obtain search warrants in certain areas from public prosecutors. By presenting the prosecutors with the software's 'hot spots' for the following days, Perrot allegedly convinced them that it was necessary to act in line with judicial policing, and therefore to issue special authorisations for identity and vehicle checks in the zones in question. This constitutes a serious violation of administrative police duties.

**Evaluation of results**

No official assessment is available, and our CADA requests to the Ministry about PAVED remain unanswered. However, there are a number of relevant academic works. The aforementioned report by the *Institut d'Aménagement et d'Urbanisme de la Région Île-de-France* drew on interviews with officers to examine the effects of PAVED.[66] Another study, by a team of researchers specialising in management who partnered with the *Gendarmerie*, examined the use of PAVED 'in the field' by 14 officers at several sites over the course of the 18-month trial period. The researchers had direct access to officers using the system and conducted regular interviews with them during this period.[67]

The authors conclude that PAVED partially meets expectations for better planning support and optimising patrol dispatch, but fell short on apprehending people committing crimes. The officers interviewed shared mixed opinions about the system's usefulness. It was not so much the algorithm's 'predictive' function they found useful, but rather its support with streamlining management.

These conclusions echo those of the researcher Camille Gosselin:

> According to its most severe critics, the tool's predictive powers are limited (compared to some of the software used by American police forces), since it is essentially based on past data and does not take sufficient account of other variables. Its analysis is therefore based on probability calculations. The platform is mainly used by the hierarchy and geared towards the chain of command. However, while it is not generally considered a finished tool, it has created expectations, particularly at the operational level. The platform aims not so much at understanding and analysing crime phenomena, as meeting the challenges of optimis-

66. Camille Gosselin. "La police prédictive : enjeux soulevés par l'usage des algorithmes prédictifs en matière de sécurité publique". Paris : IAU Île-de-France, 2019. Available at: *https://www.institutparisregion.fr/fileadmin/NewEtudes/Etude_1797/Etude_Police_Predictive_V5.pdf*

67. Cécile Godé, Sébastien Brion, and Amélie Bohas, "The Affordance-Actualization process in a Predictive Policing Context: insights from the French Military Police," in European Conference on Information Systems (ECIS) (Marrakech, Morocco, 2020). Available at: https://hal.archives-ouvertes.fr/hal-02500125. The complete study, in French, is referenced, but the Gendarmerie objected to its public release. See: *https://ideas.repec.org/p/hal/wpaper/hal-02500382.html*

> ing resources and ensuring a proper matching between crime and the territorial distribution of *Gendarmerie* manpower. At this stage, it would be premature to consider this approach as a genuine paradigm shift. While it does raise a number of questions, it does tend to reinforce the dynamics observed over the last few decades within the security forces (professional orientations, management, policies based on target objectives, etc.).[68]

Colonel Patrick Perrot is currently the *Gendarmerie Nationale*'s AI coordinator. According to sources, the *Gendarmerie* is still conducting research and development into 'predictive' policing. If these technologies prove effective, and the context favourable, they will likely be used again.

### 2.3.5 Smart Police (Edicia, municipal police forces)

Smart Police is a software platform developed by French company Edicia that aims to simplify the work of police officers. As early as 2014, Edicia announced that it had sold the software to nearly 470 municipalities across the country. It is worth noting that in 2020, only 3,681 municipalities had a municipal police service, out of the 34,839 municipalities in France at the time.[69]

Edicia was created in 2013 with headquarters in Nantes. That year, Vincent Loubert, a former Capgemini consultant, bought a software company called Access with the support of the Newfund investment fund.[70] For roughly ten years, this small company, originally created by a retired policeman, sought to develop an app for simplifying police work.

In 2019, following its rapid expansion in France, Edicia expanded internationally to the US, selling its software to local police forces, notably the Denver (Colorado) police. The company even set up offices there, with around 30 staff members.[71] The startup's HQ in Nantes grew, employing 40 new people that year, then allegedly turning a profit for the first time since its creation.

68. Camille Gosselin, "La police prédictive : enjeux soulevés par l'usage des algorithmes prédictifs en matière de sécurité publique" (Paris: IAU Île-de-France, 2019). Available at: *https://www.institutparisregion.fr/fileadmin/NewEtudes/Etude_1797/Etude_Police_Predictive_V5.pdf*, 25.
69. It should be noted that, on its website, Edicia also boasts that some of its customers are from the French Ministry of the Interior, but our CADA requests to the Ministry concerning these collaborations were unsuccessful—the Ministry claimed it had no association with Edicia.
70. Colas des Francs, Ophélie. "Le deal de la semaine : Edicia, pépite de la sécurité urbaine, lève 1,5 million d'euros". Entrepreneurs, November 27, 2013. Available at: https://entrepreneurs.lesechos.fr/creation-entreprise/aides-financements/le-deal-de-la-semaine-edicia-pepite-de-la-securite-urbaine-leve-15-million-deuros-1996304.
71. Agence API, "Le modèle éco d'Edicia bascule vers le contrat Saas", October 30, 2019. Available at: https://agence-api.ouest-france.fr/article/le-modele-eco-dedicia-bascule-vers-le-contrat-saas.

As of January 2024, Loubert had invested over €10 million in research and development over the previous 10 years. Whereas previously Edicia sold licences as a service provider, it is now transitioning to a service company model, with Edicia's customers' data stored on a company server.[72]

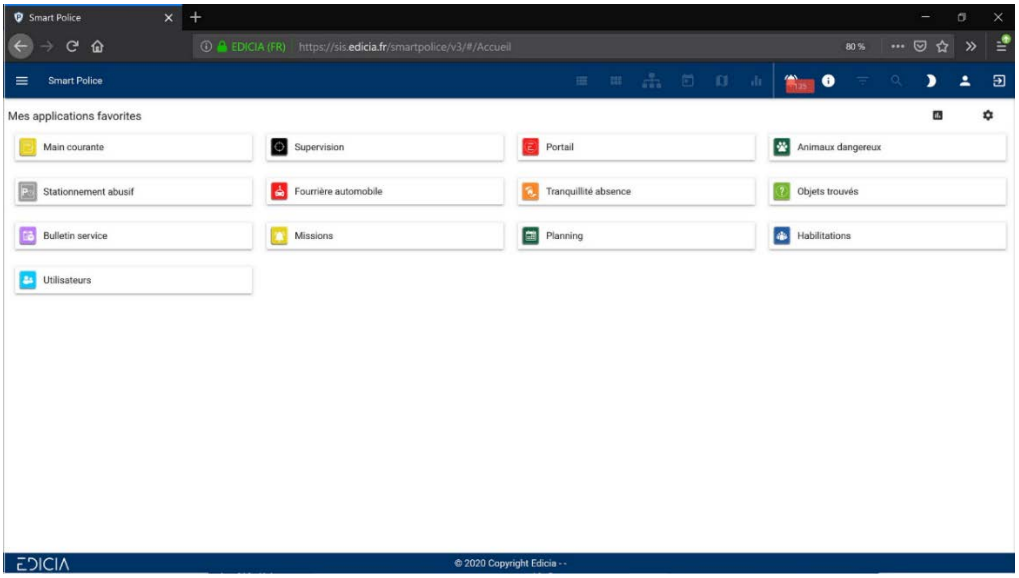



*Figure 13: Customised Smart Police home page (this version does not include the predictive module) (Smart Police user manual).*

**Overview of Smart Police**

As already explained (see section 1.3.3), Smart Police is a software platform. It was not originally created with a 'predictive' function. The platform allows police officers to write reports 'in the field' using their phones or tablets, add photos to a database, report events or draw up official reports. The platform also allows commanders to monitor field teams from their offices, map incidents, consult their reports and receive various statistical indicators in real time. They can also view photos taken during an intervention (e.g. during a demonstration, see figure 16).

A press article from 2018 explained how the software is used:

> Maurepas, Wednesday May 30, 10 a.m. Olivier Pruvost, head of the municipal police force, checks the central station screen. A burglary took place the previous evening. A few dots flash on the screen: they symbolise the geolocated police officers on patrol in town. Five hundred metres

72. See the RGPD compliance certificate issued by the law firm Olivier Iteanu, obtained via a CADA request to the city of Libourne: https://cloud.laquadrature.net/apps/files/?dir=/LQDN/Technopolice%20Data/Libourne/Edicia&openfile=994213.

> away, Hamid is running an OTA, an "*Opération Tranquillité Absence*." He walks around a house which the owners, who are away for a few days, have placed under the vigilant care of the local police. Two clicks on his smartphone and he's filled in his card, nothing to report (…) Two years ago, the towns of Maurepas and Coignières went digital: the town's eleven police officers were given phones equipped with a range of professional software, which they turned into their personal desks. Now, officers draw up reports in the field, adding photos to the report, which they send directly to headquarters. No more tedious trips back and forth to the office (…) In Maurepas, police officers have also abandoned the good old ticket book, used to write up parking tickets. The municipal policeman writes his ticket four times faster—30 seconds instead of two minutes. Addresses are filled in almost automatically, and he scans the licence plate directly and photographs the offence, if necessary, to keep a record.[73]

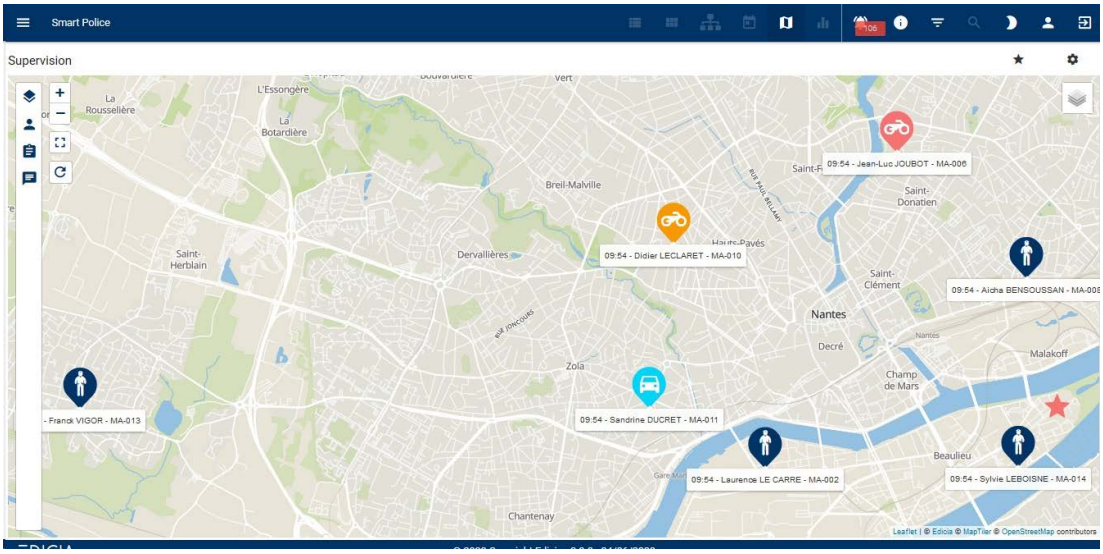



*Figure 14: Customised home page for the 'Supervision' module, displaying the geographical distribution of teams (patrol car, officers on foot, two-wheel vehicles) (Smart Police user manual).*

73. Delphine Dechaux, "Cet incroyable éditeur nantais de logiciels qui améliore la sécurité des villes en numérisant la police municipale et la RATP," Challenges, June 1, 2018. Available at: *https://www.challenges.fr/high-tech/cet-incroyable-editeur-nantais-de-logiciels-qui-ameliore-la-securite-des-villes-en-numerisant-la-police-municipale-et-la-ratp_590969*

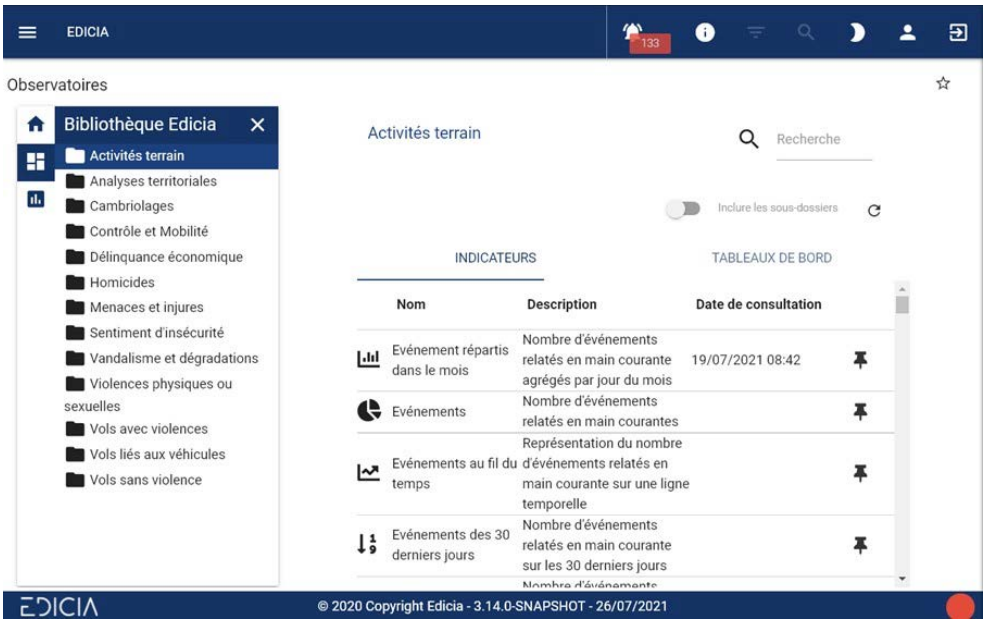



*Figure 15: List of indicators available in the 'Observatory' module and, on the right, a list of offences for which indicators can be displayed (Smart Police user manual).*

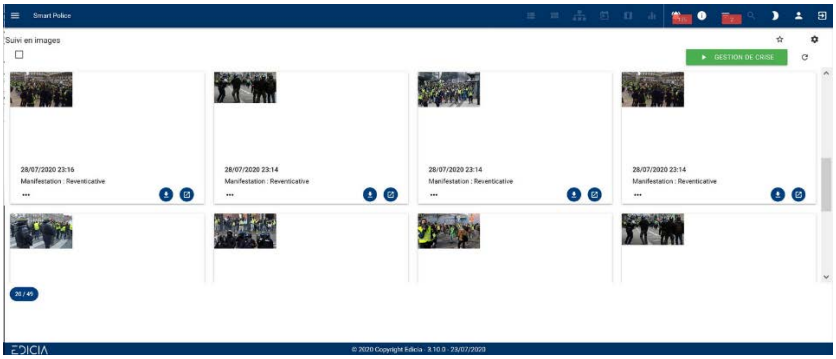



*Figure 16: Images from the 'Image Tracking" menu, which displays thumbnails of the latest photos taken by officers via the 'Complaints' menu. In the example shown, the 'Tracking details' view reveals a photo taken during a Yellow Vests demonstration (Smart Police user manual).*

## Towards 'prediction'

As early as 2014, Edicia promoted the evolution of its Smart Police software by talking about the next step: 'prediction'. The company proposed nothing less than the ability to issue warnings such as: "Be careful, tomorrow, in such and such a place, there is such and such a risk." In an interview from February 2014,[74] company representatives were talking about risk models Edicia was working on. These were being developed in partnership with the LINA (*Laboratoire d'Informatique de l'Université Nantes Atlantique*, the computer science laboratory at Nantes Atlantique University). Edicia's marketing director said:

> Our tool enables us to gather data extracted from rumours on social media platforms, to qualify this data and then to send out resources, with the tool also managing the availability of agents. We process this data in a risk assessor. For example, if there's a rumour of a giant party, it could bring together 10,000 people if the weather's fine, but only 2,000 if it's not. If they have the information, the security officers will be able to deploy the appropriate resources.

It wasn't until 2018 that the company filed a patent for this 'predictive' component of their software, entitled *Procédé et système de surveillance et de prévention de dysfonctionnements en sécurité territoriale*.[75] The platform's 'predictive' module was developed through a doctoral thesis funded by the French Ministry of Defence's *Direction Générale de l'Armement*. The patent describes the system as follows:

> A system for monitoring and preventing dysfunction in a territory divided into a plurality of interconnected areas, each comprising elements to be monitored; each area being equipped with a plurality of sensors distributed in each area and adapted to provide measurement signals representative of the violation of a property of the elements to be monitored according to predetermined rules.

Once again, on the basis of situational crime prevention, Edicia's software aims to 'predict' events that are likely to occur in a given area—ideally both 'regular' and 'rare' events. Based on the idea that "risk is a function of the probability, severity and detectability of a risky event," the system sets out to "produce" predictions from a number of different signals.

74. Maryse Gros, "Edicia associe big data et sécurité urbaine", Le Monde Informatique, February 19, 2014. Available at: *https://www.lemondeinformatique.fr/actualites/lire-edicia-associe-big-data-et-securite-urbaine-56624.html*
75. Annie Bourget and Kante Thierno, "Procédé et système de surveillance et de prévention de dysfonctionnement en sécurité territoriale", Institut national de la propriété intellectuelle, patent no. 3079952 (Courbevoie, 2018), *https://cloud.laquadrature.net/s/kTxZtfxqmyGYP32*

The exact nature of these signals is not specified. The patent is vague on the precise functioning of the system. All that is mentioned is the visualisation of existing "prevention barriers" (vehicles, patrols, alarms, etc.) and "action means," which, in the event of a risk in a given area, make it possible to check whether foot patrols, vehicles and other "prevention barriers" are in line with the nature and degree of the identified risks. The patent also mentions urban planning data (yet again echoing the RTM approach), environmental and weather data, future national and local events, socio-demographic and electoral data, and more.

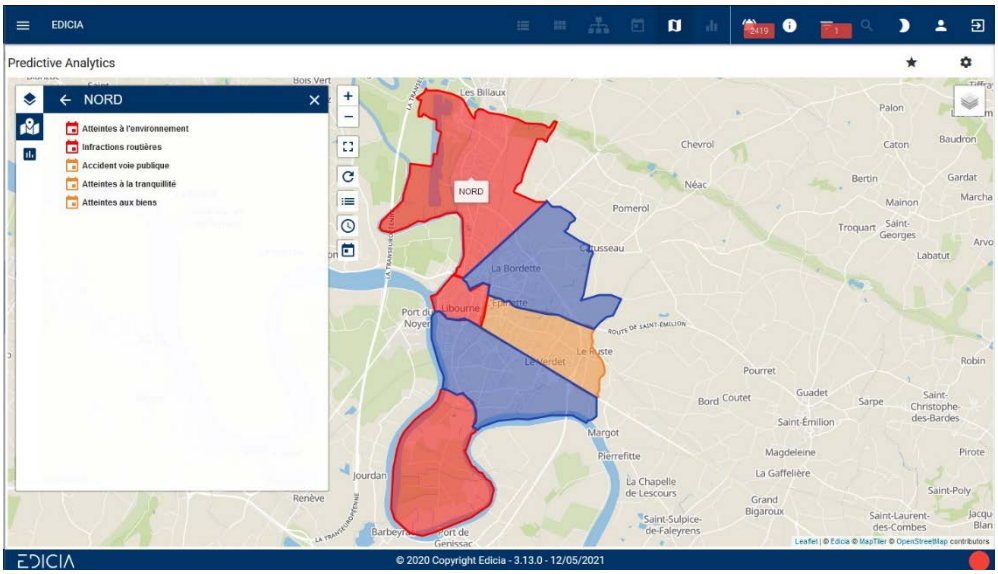



*Figure 17: A view of Edicia's 'predictive' module (Smart Police user manual).*

As input data, Edicia plans to combine two data sources:

- a 'field' database, containing information entered into the Smart Police software:
  - reports and official statements on offences;
  - geolocation of officers and their vehicles, etc.; and
- an 'external' database:
  - urban planning;
  - socio-economic profile of the population;
  - weather forecasts; and
  - "rumours, public information from social media and security partners: high school principals, social housing landlords," among others.

These fields are encouraging the illegal recording of data, with the risk of amplifying discriminatory policing practices.

Cities such as Nice, Libourne, La Rochelle, Charleville-Mézières, Marseille and Le Pré-Saint-Gervais are cited by various sources as Edicia customers. Despite submitting 15 CADA requests, we were unable to obtain corroborating information on the real uses of Smart Police, particularly with regard to its 'predictive' component. Our CADA requests only received a response on the city of Nice's use of the 'predictive' component. The response implied that it had never been used. This contradicts information received from elsewhere, which mentioned workshops organised by Edicia with municipal police forces to refine the 'predictive' model.

The documents provided by Libourne also include no mention of the use of the 'predictive' module. This is despite boasts on Edicia's website about results achieved, and an active collaboration between the *Police Municipale* and the *Gendarmerie Nationale*:

> Thanks to the predictive algorithms integrated into the platform, the city of Libourne provides its partner, the *Gendarmerie Nationale*, with a daily risk map of the area for the next 24 hours.

Bourges, a town with 65,000 inhabitants and 45 *Police Municipale* officers, paid €18,000 to access Smart Police's basic functions for just 18 months.[76] It can therefore be assumed that the 'predictive' module is relatively expensive. The financial cost can be explained by the fact that city-wide implementation requires significant work on input data and parameters. The high cost could explain the limited uptake of this technology, as well as plausible disillusionment of regions testing it. They may have been disappointed by the gap between Edicia's dazzling promises, and the system's actual impact. This appears to be a common disillusionment among police forces.

However, the Smart Police user manual reveals a number of concrete details on how the 'predictive' module works.[77] This feature, named 'predict', is accessible via the software's 'command post' and therefore only by commanding officers. The manual says little about how the model works in practice, or how it derives information from the data compiled in the software:

> The Predictive Analytics engine recovers data from the "Complaints" module, the "Active Vigilance" module, manpower deployment, as well as data from connected third-party applications (social media, etc.).[78] Continuous improvements in the field of artificial intelligence mean that environmental data (traffic, weather conditions) can be linked to operational data from agents in the field.

76. We uncovered this information in a document obtained in response to a CADA request.
77. The Smart Police manual is available in HTML format at: *https://technopolice.fr/police-predictive/manuel-edicia/Edicia.html*
78. According to some of our sources, and contrary to what the manual seems to indicate, the inclusion of social media in Smart Police input data was still in the planning stage in 2022. We were unable to obtain further information on this crucial point for civil liberties.

According to the manual, risk 'predictions' are tailored to the needs of the customer, taking into account three parameters:

- the city's calculation model;
- the calculation input parameters (e.g date range);
- facts and resources by neighbourhood and time of day.

When the 'predictive' module is launched, it displays a map of the city. This has a colour gradient according to risk levels per neighbourhood, and per six-hour period, over the next 24 or 48 hours. Red correlates to high risk, orange to medium risk and blue to no risk. The user can superimpose information on planned patrols, and adjust the schedule according to the danger scores generated by the algorithm.

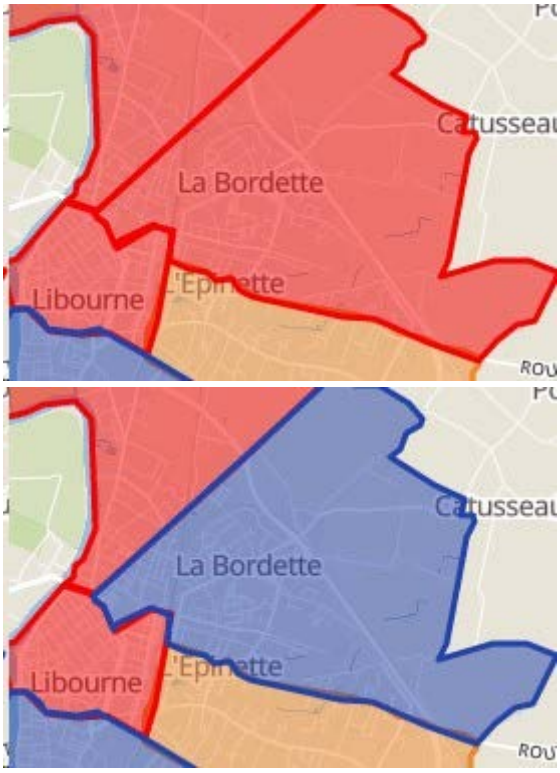



*Figure 18: When users de-select the risk type ('traffic violations' in this example), the system displays the general risk level of the La Bordette neighbourhood without including the estimated risk level for traffic violations. The system then takes on the colour of the highest risk level determined for the active 'risk type' (in this example, it changes from a high to a low risk level) (Smart Police user manual).*

### Edicia AIs

The 'predictive' module developed by Edicia allegedly uses two AI techniques to automate its predictions:

- a 'symbolic' AI, based on criminological assumptions coded according to 'decision trees'; and
- a 'connectionist' AI, based on machine learning with iterative correction of automatically generated results.

This combinatorial method was presented as being particularly innovative in academic papers by one of Edicia's data scientists.[79]

The machine learning model is trained on a large quantity of data relating to the same type of event. This type of algorithm can therefore only correctly take into account events recurring relatively often. The model requires another type of algorithm for rare events, known as 'symbolic' events. For this, the police provide Edicia's data scientists with their first-hand knowledge: workshops are regularly organised between Edicia's data scientists and representatives of the police force.

According to sources we spoke to, this 'feedback' has led to the introduction of dubious knowledge, such as "the fact that petty crime encourages more serious crime." This statement that evokes the theory of 'broken windows policing' (examined in more detail in section 3.3). This also runs the risk of 'whitewashing' discriminatory logics by "putting them into algorithms," giving them a veneer of objectivity.[80]

Edicia's basic 'predictive' model, whose parameters are unknown, is also based on 300 'general categories'. These can be adjusted to customise Smart Police predictions and metrics according to customer needs and feedback. In any case, R&D on Edicia's 'predictive' module seems to be continuing, with the aim of improving the integration of sources from social media and adding data from victim surveys.

79. Thierno Kante and Philippe Leray, "A Probabilistic Relational Model Approach for Fault Tree Modeling," in Advances in Artificial Intelligence: From Theory to Practice, ed. by Salem Benferhat, Karim Tabia, and Moonis Ali, Lecture Notes in Computer Science (Cham: Springer International Publishing, 2017), 154 62.
80. The "broken windows" theory posits that "minor forms of public disorder, such as graffiti, littering, begging and prostitution, if left unaddressed, will lead to neighbourhood decline and an increase in serious criminal activity." See: Bernard E. Harcourt, Illusion of Order: The False Promise of Broken Windows Policing (Harvard University Press, 2005), 3.

## 2.3.6 M-Pulse (Marseille municipal police, Engie Solutions)

In November 2017, Marseille city council publicly announced the creation of the *Observatoire du Big Data de la Tranquillité publique* (Big Data Observatory for Public Safety). Marseille's vast geographical area has a total of around one million inhabitants.[81] Under the terms of a partnership with the Bouches-du-Rhône prefecture, Engie Inéo (now Engie Solutions) has developed a data-sharing platform called M-Pulse. This has an online interface for municipal and national security services.[82] Following the coronavirus lockdown in March 2020 and municipal leadership changes in June 2020, the project continued under a new political narrative. The general public was invited into consultations on the interface, albeit for unclear purposes.

### Initial objectives

Through CADA requests, we obtained several documents that shed light on M-Pulse's objectives. The '*Cahier des clauses techniques et particulières*' (CCTP), the specifications for companies responding to the city council's call for proposals, provided a wealth of information on the project's ambitions.[83] The document summarises the project's objectives, outlines the range of data sources to be integrated into the platform, and defines its role: to study the past and shed light on the present, to anticipate future events.

The *Observatoire* was thus presented as a highly ambitious experiment in centralising and pooling urban data, to better understand what's going on in the city, with a direct operational objective. M-Pulse advertised itself as a large integrated technology platform "based on big data and machine learning methods."

The overall ambition of the project was to "break down silos" and share information between the various "stakeholders in public safety" in order to optimise their interventions. "To take full advantage of the knowledge of each individual and of the specialised data silos, the sharing of information is an essential prerequisite," according to the project specifications.

81. City of Marseille, "Création d'un outil Big Data de la Tranquillité Publique et prestations d'accompagnement — Cahier des Clauses Techniques Particulières (CCTP)", January 2017. Available at: *https://data.technopolice.fr/fr/entity/69yuoeous9u?page=24*
82. Ministry of the Interior and City of Marseille, "Convention de mise à disposition de données pour la plateforme Big Data de la Tranquilité Publique," January 13, 2020. Available at: *https://data.technopolice.fr/fr/entity/eaxer6u3c5?page=2*
83. City of Marseille, "Création d'un outil Big Data de la Tranquillité Publique et prestations d'accompagnement — Cahier des Clauses Techniques Particulières (CCTP)".

There is explicit mention of the project's 'predictive' dimension. The platform is supposed to "analyse what happened (yesterday)", "assess the current situation (today)" and "anticipate the future or probable situation".[84] For all these tasks, three main use cases were identified:

- "Analysis and forecasting of criminality and public safety problems in public space";
- "Analysis and forecasting of problems related to the use of public space";
- "Analysis and forecasting of traffic flow, parking and road safety issues."

The initial data sources were varied, with both "raw" and more or less "structured" data supplied by the city of Marseille's *Délégation Générale à la Sécurité* (DGSEC). These included intervention reports and the database of fines issued by the *Police Municipale*. Data was also supposed to come from other public institutions such as hospitals, the Marseille transport network (RTM), the city's marine fire department, the port authority and weather services. Data from private institutions, such as road traffic surveillance images (including those from private motorway), and data extracted from social media were also included.

Particularly concerning is the document's reference to:

> ...assessment of the risk of dangerous gatherings by analysing tweets, based on the identification of actors (Who is speaking? Who is acting? Who is interacting with whom?) and feedback from conversation threads (Who is organising? Who is the first to post?).

The project specifications include amongst the city's external data partners as the Ministry of the Interior's databases and private companies. While it appears that no private partner actually contributed data in the end, M-Pulse's designers included partnerships with telecom operators in their original "grand vision." This would, for example, produce statistics and 'heat maps' showing the geographical distribution of the population using mobile phone location data.

The budget €1.5 million is relatively modest compared with the €40 million allocated to the installation of 500 new surveillance cameras in the city. It is largely covered by a €600,000 grant from the European Union's European Regional Develop Fund programme, and a €600,000 grant from the Bouches-du-Rhône regional authority. A third of this budget was used to purchase storage space on the Big Data Appliance servers operated by the company Oracle.[85]

84. Ibid, page 12.
85. "Témoignage Ville de Marseille : Projet Big Data de la Tranquillité Publique avec Oracle Big

**Limited implementation**

Ultimately, the solution implemented from 2019 in Marseille consists of two components: a software interface (API) to integrate the data into the system, and a web interface to access it. These had a total cost of €1 million.

There are three use cases outlined in the specifications, as noted above. The first, "analysis and forecasting of criminality and public safety problems in public space," is undoubtedly the most concerning from a human rights perspective, due to its focus on surveillance and profiling. The aim of this use case is to adapt police deployments on the basis of 'predictions'. The aim is to prevent situations of obstruction and danger when scheduled events are too close together in space or time, or due to other factors (weather, road traffic, simultaneity of events).

The potential legal implications of this first use case may be the reason why the pilot focused solely on the second use case: "analysis and forecasting of problems linked to the use of public space." Other reasons the 'predictive' use case was not implemented may be a desire to simplify implementation and to improve the social acceptability of the project.

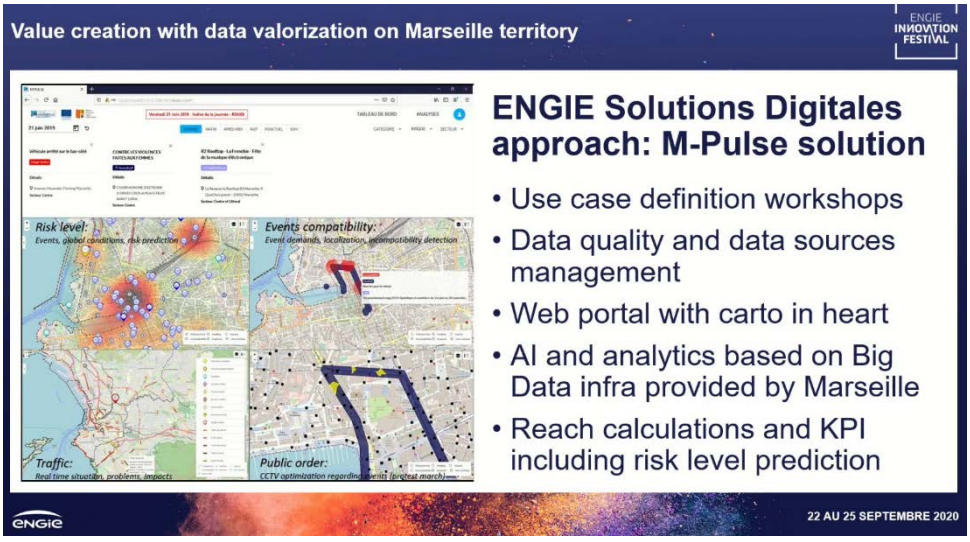



*Figure 19: A slide presented at Engie Ineo's internal seminar on the M-Pulse project (source: 'Jumeau numérique" presentation).*

Data", 2017, available at: *https://www.youtube.com/watch?v=qYmbfQ-8BB4*

The M-Pulse web interface is a "two-dimensional map portal." The input data is entered into a natural language processing algorithm, which extracts the relevant information from the municipal databases and places it on a map of the city. The map includes:

- religious, cultural and sporting events;
- demonstrations;
- markets;
- construction sites;
- road accidents;
- environmental damage;
- traffic disruptions;
- illegal parking;
- film shoots;
- damage to property;
- public order disturbances; and
- police operations.

For the safety in public space use case, stakeholders identified questions for the 'predictive' technologies to answer:

- Where are known events located?
- Is there any incompatibility between known events and their impact on public space?
- How can we identify which events need special attention?
- What is the video coverage rate of a geographical area?86

The second step was to identify the relevant data sources. Several databases within the city's jurisdiction provided M-Pulse with over 3.5 million data points. The databases used for this step were:

- the PATIO database, which lists all tourist and cultural events in the Bouches-du-Rhône prefecture, supplied by 70 bodies in the region (mainly tourist offices);
- the database of municipal bylaws: planned roadworks, film shoots, building permits, etc.;
- the Police Municipale databases, which we know contain complaints, as well as data from the iPolice app87 used by municipal officers in their field missions (containing mission orders, official reports of tickets and misdemeanours);88
- the database linked to the geographical location of video surveillance cameras.

86. Ville de Marseille, "État des lieux du projet d'Observatoire Big Data de la Tranquillité Publique," October 3, 2019, available at: https://data.technopolice.fr/fr/entity/litu04kbx4e?page=2; Ville de Marseille, "Dossier de présentation — Projet Big Data de la Tranquillité Publique," October 7, 2019, https://data.technopolice.fr/fr/entity/b6mmkge146?page=1.
87. This is probably a version of Edicia's Smart Police software, since Marseille is a customer of the company.
88. Our CADA requests of 2022 concerning the use of Smart Police by the city of Marseille have remained unanswered.

Another crucial data source for the platform is the Ministry of the Interior. Many public events are registered by the Ministry via the prefectures. Through a CADA request, we were able to obtain the data-sharing agreement between Marseille City Council and the Ministry.[89] According to the agreement, there are three categories of data shared between the Bouches-du-Rhône prefecture and municipal authorities:

- announced events;
- planned police operations; and
- "official visits" (government members, foreign diplomats and "sensitive public figures").

According to a document signed by the city's director of legal affairs (obtained through a CADA request), the prediction algorithm is a 'random forest' model. In her words:

> The algorithm used for the learning model is based on decision trees. These trees are constructed using the "random forests" algorithm. Once similar past events have been interpolated, the facts observed in the vicinity of these past events are "transferred" to the predicted event, or more precisely, the score computed for each past event from the combination of observed facts is transferred to the predicted event. Lastly, we are still unable to provide you with an impact assessment because, as already indicated to you in October 2019, this processing does not present a high risk for the rights and freedoms of individuals and does not fall within the list of types of processing for which the CNIL has deemed it mandatory to carry out an impact assessment.[90]

**Even more limited use**

At the end of 2019, as per their contract with Engie Ineo, the city of Marseille was supposed to receive a report with "quantifiable results" on the progress made in managing public spaces thanks to M-Pulse. Allegedly, the platform would also incorporate new datasets, including: weather forecasts, transport data, events organised on social media such as Facebook, and population distribution data derived from mobile phone locations.[91]

The project fell behind schedule before the COVID pandemic put a definitive stop to the use of M-Pulse. With repeated lockdowns, there were no

89. Ministry of the Interior and City of Marseille, "Convention de mise à disposition de données pour la plateforme Big Data de la Tranquilité Publique."
90. Marie-Sylviane Dole, "Demande d'accès aux documents publics relatifs à la surveillance algorithmique de l'espace public marseillais", Direction générale adjointe de l'action juridique, City of Marseille, July 6, 2020. Available at:https://data.technopolice.fr/fr/entity/yjhsdstl1a?page=1.
91. Jamal Al Hassani, "À Marseille, une carte dopée aux données pour gérer l'espace public", Journal du Net, August 29, 2019. Available at: https://www.journaldunet.com/economie/services/1443369-marseille-carte-plateforme-big-data/.

public events for months. The *Police Municipale* concentrated on other tasks, subsequently abandoning a technology that had become obsolete. When events in public spaces recommenced, the municipal police apparently failed to reintegrate M-Pulse into their daily operations.

A source within the *Police Municipale* suggested that the platform was currently under-utilised. The source outlined the issue of information silos, which the system was meant to overcome: "there are departments that don't report [information], that don't contribute to this tool." This is despite the fact that the platform's stated aim was to overcome the compartmentalisation of services. Apparently, the *Police Nationale* was not transmitting the information it was supposed to provide regularly. For these reasons and more, the source admits that the initial enthusiasm surrounding M-pulse quickly waned: "We're not using it as much as we should."

### 2023: Re-engineering the project as a mapping app for citizens

Initially, it appeared that the change in municipal council in June 2020 would lead to the project being put on hold. The new left-wing majority had pledged during their election campaign to introduce a moratorium on police surveillance technologies.[92] The new municipal authorities declined most of our requests for interviews about M-Pulse. Nor did they respond to our written questions or CADA requests, despite the favourable opinion of the *Commission d'Accès aux Documents Administratifs* to our requests.

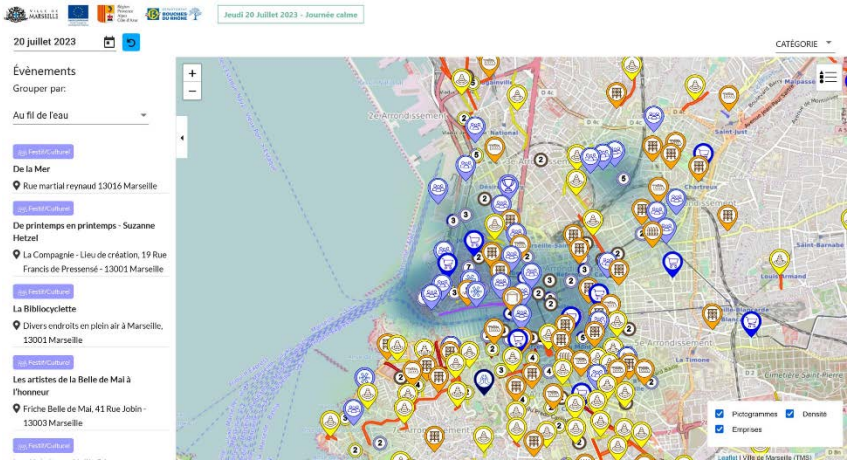



*Figure 20: Screenshot from the public version of the M-Pulse website.*

92. La Quadrature du Net, "Le Printemps Marseillais dans la Technopolice ? Au fond à droite", June 8, 2021. Available at: https://www.laquadrature.net/2021/06/08/le-printemps-marseillais-dans-la-technopolice-au-fond-a-droite/.

It was not until June 2023 that more information became available. City councillor Christophe Hugon from the Pirate Party[93] presented the 'new version' of M-Pulse to the public at the annual conference of the OpenStreetMap community in France,[94] which was held in Marseille that year.

Hugon's presentation offered a narrative that he had been crafting for months.[95] On entering the city council after the June 2020 municipal elections, Hugon asked to see the platform that was "causing a stir on the outside". He felt the previous majority had presented "anxiety-provoking" information, "like analysing social media to find out who's doing what, where, etc." To his great surprise, M-Pulse appeared nothing more than "simply a public service management project, because that's what it's always been about in the end."

Subsequently, M-Pulse's interface was made accessible to the public, stripped of any security, policing or predictive dimensions.[96] Why would the local authorities allow public access to this application? According to Hugon, M-Pulse would provide value to the city's residents:

> We think that having a map that allows the people of Marseille to know where the markets are, where the events are, to know where there will be a lot of people... if you want to live in your city and go where there are people, it's interesting information.

He also uses the example of tourists who would like to "visit the markets," and could then view the market icons to know where to go. It is hard to see how this app differs from the services offered by Google Maps. Hugon remained vague on the other functions. He admitted that some useful functions have been removed for the public version, such as the display of police presence in real time, for "obvious" security reasons.

But if Marseille's municipal police force is still using M-Pulse, what value do they see in it? What about its 'predictive' functions? The data sources it uses? The use of social media, and so on? When questioned at the end of his presentation, Hugon simply avoided the subject. He said the uses devised by the previous council were, for the most part, illegal and unrealistic, and have been abandoned.

93. Hugon is a member of the new municipal majority and delegate to the mayor for open data and digital responsibility.
94. "State of the Map France 2023".
95. Ibid.
96. Available at: *https://m-pulse.marseille.fr/map*

The mayor's office refused to comment on how the police actually use the software. Nor will it amend the original contract in light of the system's change in purpose, wrongly claiming that it is impossible to change a contract that has already been signed. Nonetheless, as it stands, in terms of civil liberties, M-Pulse does not appear to involve the same concerning levels of surveillance and 'prediction' as PAVED, Smart Police and Risk Terrain Modelling.

# 3. Critical analysis of these systems

In this section, we present a critical analysis of the 'predictive' policing systems studied in this report.

## 3.1 Potentially discriminatory parameters

The first and most serious issue with these systems is the likelihood that they will lead to discrimination, resulting in significant and serious policing and legal consequences for the marginalised groups targeted.

Police data in these models will always represent the structural biases in policing. However, the police also use socio-demographic data as part of these 'predictive' models. This means information on peoples' backgrounds or where they live is being used to 'predict' crime and criminality.

For example, the PAVED 'predictive' model uses 15 socio-demographic parameters. According to the developers, these are strongly correlated with crime. However, there is no transparency around these 15 variables, let alone any attempt to demonstrate a true cause-and-effect relationship. Similarly, PredVol was based on over 600 socio-demographic parameters with no indication of how the software used them. The same can be said for the variables used by Edicia's Smart Police software, though in this case we have even less information about the exact nature of the selected variables.

The use of these socio-demographic variables may cause discrimination, as was the case with the algorithms used by the *Caisse Nationale des Allocations Familiales* (CNAF).[97] For example, high levels of unemployment, poverty, or a high proportion of people born outside the European Union in the neighbourhood in question may lead to higher 'predicted' risk scores.

97. La Quadrature du Net, "Notation des allocataires : l'indécence des pratiques de la CAF désormais indéniable," November 27, 2023. Available at: *https://www.laquadrature.net/2023/11/27/notation-des-allocataires-lindecence-des-pratiques-de-la-caf-desormais-indeniable/*

This is especially likely for the PAVED system. The input data used to establish 'predictions' includes nationality and immigration details, household income and composition, and level of education. All these indicators run the risk of targeting marginalised groups and those most impacted by structural racism. It may also kead to feedback loops, where those same groups are repeatedly targeted (discussed further below).

Concerns about deducing risk scores from discriminatory variables is increased by the "black box" nature of these systems, given their reliance on machine learning techniques.[98] A system is said to be a 'black box' when it performs complex calculations which it does not clearly explain, or which cannot be understood by someone who wants to know how the output, such as a risk score, was produced. Some crucial questions may remain unanswered: which input variables contributed most to the production of the output? What exact characteristics did the model derive from the input data? How does the output change the input variables? It is unclear whether the designers of 'predictive' policing systems using machine learning techniques would be able to answer these crucial questions.

## 3.2 Feedback loops

The criticism is well-known, but worth repeating: the use of 'predictive' policing systems raises major risks of reinforcing the historic and targeted policing of neighbourhoods where racialised and low-income people live. For example, they may face increased police surveillance, identity checks, and use of coercive powers.

The use of such 'predictive' systems, based off and using historic police data, lead to certain areas being designated as 'high risk'. These are often areas which have historically been over-policed and targeted, and are therefore over-represented in police data. When a large number of police patrols are sent to these areas, in response to the system's 'predictions', they are more likely to observe or find crime or criminality (even minor offences), leading to more data on those areas being collected. This in turn, will be fed back into the 'predictive' system. increasing the probability that the area will be perceived as 'high risk' in the future.

98. Will Knight, "The Dark Secret at the Heart of AI." MIT Technology Review, April 11, 2017. Available at: *https://www.technologyreview.com/2017/04/11/5113/the-dark-secret-at-the-heart-of-ai/*

In this way, 'predictive' policing algorithms produce a self-fulfilling prophecy, or feedback loop, by concentrating police operations in areas already impacted by discrimination and targeted policing. Once again, this perpetuates long-standing trends, highlighted by sociologists since the 1960s, that crime statistics tend to reflect police activity rather than actual trends in crime.[99]

While more research is needed to fully demonstrate this, one of our interviewees claims to have observed this feedback loop in operation in marginalised neighbourhoods. That is, there was an increase in recorded crime in areas already exposed to structural discrimination.

As far as we know, the risks of feedback loops have not been meaningfully considered by the designers of these systems. In fact, these risks have been dismissed entirely. According to some proponents of 'predictive' policing, the aim is not to increase the number of offenders caught in the act, but to pre-empt crime by increasing the presence of police forces in sensitive areas. In short, their intention is to 'prevent' crime and, above all, increase the population's sense of security through a stronger police presence.

Whatever the stated intention, increased police presence in specific areas will undoubtedly lead to increased criminalisation and police violence in those areas. This is especially so in neighbourhoods where relations between the police and the local community are notoriously poor.

## 3.3 False criminological assumptions

Another danger associated with these systems is that they entrench problematic criminological premises in police operations. This is further amplified by the lack of transparency surrounding them.

As previously mentioned, proponents of 'predictive' policing refuse to engage in the general understanding and social analysis of crime. They include no mention of the social causes of crime, such as economic insecurity, exclusion and discrimination, or of the social violence of public policy. When 'predictive' policing models attempt to use explanatory approaches in their algorithms, they often rely on "knowledge" that is of dubious relevance.

99. Biderman, Albert D., and Albert J. Reiss. "On Exploring the 'Dark Figure' of Crime". The Annals of the American Academy of Political and Social Science 374, No. 1 (November 1967): 1 15.

Some of these problematic premises appear in research articles by PAVED's main developer, Colonel Perrot. In these articles, he proposes basic hypotheses about crime (e.g., crime as a "constantly evolving phenomenon"). He refers to "weak signals" and other "warning signs", the scientific nature of which is highly disputed. Similarly, Edicia's 'predictive' module is based on the notion that crime has a geographical spill-over (or "contagion") effect. The company alleges that "petty criminality leads to more serious criminality."

These systems evoke some of the most dangerous trends in criminology, such as the 'broken window theory'. Developed in the US in 1982 by James Wilson and George Kelling, this theory was used during the Reagan era to redirect police operations to focus on behaviour and minor crimes. Using a psychological study from 1969, Wilson and Kelling produced a ready-made model for the conservative regime. In a free interpretation of the expression "he who steals a pin will steal a greater thing," they asserted that repressive action on small offences would help to curb larger crimes.

The broken window theory has been largely discredited as not having an appreciable impact on crime.[100] It masked the disastrous consequences of neoliberal social policies pursued at the time, and criminalised minor crimes in an effort to criminalise the poor. Today, this problematic theory is adopted and implemented by these data-driven policing systems.

## 3.4 Possible abuses of power

We were unable to find information on the specific instructions given to police officers patrolling areas labelled as 'high risk' by predictive systems. However, a well-informed source told us that, because of PAVED, the public prosecutor authorised the *Gendarmerie* to place patrolling officers in transit areas and stop nearby vehicles. This involved checking drivers' licence plates and licences, and in some cases searching vehicles.

If accurate, this would mean that preventative checks were carried out on the sole basis of a technology founded on dubious premises, the effectiveness of which has never been evaluated. This demonstrates the gross disproportionality of these liberty-restricting measures, and which is all the more questionable given that preventative controls create a strong feeling of injustice, discrimination and mistrust.[101]

100. See for example Harcourt, Illusion of Order.
101. Epp, Charles R., Steven Maynard-Moody and Donald Haider-Markel. Pulled over: How police

## 3.5 Correlation is not causation

Another fundamental issue with the use of 'predictive' policing systems is that they use statistical correlations to produce these so-called 'predictions' of crime and criminality. They do not make use of data that actually show a causal relationship between the data and the predicted acts. In fact, whether through bad faith or ideological laziness, the developers maintain a deep confusion between correlation and causation (or at least refuse to distinguish between the two).

Admittedly, correlations between socio-demographic data and crime have long been used in criminology in attempts to understand criminal behaviour and design ad hoc policing strategies.[102] In this sense, 'predictive' policing is another manifestation of a long-term trend. However, the relevance of correlations to explain or predict criminal behaviour has been widely questioned by sociologists and statisticians, who note that criminology is still unable to conclusively determine the causes of crime.[103] As a body of theory and practice, criminology also pays little heed to the risks of structural discrimination. or the systemic issue of mass incarceration to which it contributes.

It is often said that 'correlation is not causation'. We cannot deduce a cause-and-effect relationship between data variables solely on the basis of an observed association between them. The scientific approach relies significantly on the observation of correlations (two variables observed together). However, the explanatory value of these correlations must be investigated further through an empirical approach. In other words, correlations can lead to hypotheses. However, these must be tested using relevant methods to truly establish whether there is a cause-and-effect relationship.

Empirical studies have highlighted that "geographical changes have an impact on the statistical relationship between crime and socio-economic status."[104] If the links between socio-demographic characteristics and criminal behaviour differ from one region to another, national crime prediction

stops define race and citizenship. University of Chicago Press, 2014.

102. See, for example: Scott J. South and Steven F. Messner, "Crime and Demography: Multiple Linkages, Reciprocal Relations," Annual Review of Sociology 26 (2000): 83 106; see also the Wikipedia entry on correlations in criminology, available at: https://en.wikipedia.org/wiki/Correlates_of_crime.

103. See for example Greg Ridgeway, "Experiments in Criminology: Improving Our Understanding of Crime and the Criminal Justice System," Annual Review of Statistics and Its Application 6, no. 1 (2019): 37 61.

104. Peter Francis Kitchen, "Exploring the link between crime and socio-economic status in Ottawa and Saskatoon: A small-area geographical analysis" (Department of Justice Canada, Research and Statistics Division, 2006).

models (e.g. those used by PAVED) do not necessarily reflect local realities. They therefore lead to biased predictions. Again, these potential biases do not seem to be taken into account by the developers of 'predictive' policing systems.

## 3.6 The limited effectiveness of this technology

Even if these 'predictive' policing systems were to prove effective from a police perspective, they would pose major problems in terms of social justice and respect for human rights, because of concerns around discrimination. However, despite the absence of any official evaluation, the available data suggests that 'predictive' models have limited success in achieving the objectives set out by the police themselves.

Regarding their effectiveness in terms of preventing or pre-empting crime, the few existing independent studies offer results that are, at best, ambivalent. This is the case with research from two economists who have attempted to evaluate PAVED using national crime statistics.[105] Using a method they themselves describe as 'quasi-experimental', the authors conclude that car thefts have fallen in the pilot regions, possibly due to the increase in police patrols in these areas and their deterrent effect on this specific type of crime.[106]

That said, it is not certain that the recommendations produced by PAVED are directly responsible for this trend. Other factors would need to be taken into account to explain this result. At the same time, however, the rate of house burglaries did not change significantly.

What's more, these technologies appear to be far from convincing even for their users. Despite the original plans, PAVED's widespread use within the *Gendarmerie Nationale* never materialised. Following an experimental phase from 2017 to 2019, the decision was made to shelve the software. Prior to PAVED, the ambitious PredVol was also abandoned. And while M-Pulse has been revived as part of the 'citizen rebranding' pushed by Marseille's new city council majority, its security uses now seem relatively marginal.

105. Lecorps and Tissandier, "PAVED with good intentions: an evaluation of the Gendarmerie predictive policing system."
106. The reduction in the number of car thefts was estimated at between -5% and -3% per 10,000 inhabitants, i.e., an average reduction of between 114 and 68 vehicle thefts per region per year for the 11 regions where the software was tested.

Why is this? The opacity surrounding these experiments makes it impossible to say for sure, but the most likely hypothesis is that the technology adds very little real value to existing police knowledge and beliefs. The use and maintenance of these systems may also involve excessive organisational and technical complexity. In fact, feedback from use 'in the field' seems inconclusive. PredVol is no better than human deduction. PAVED proves disappointing in terms of 'predictive' capacity and does not result in increased numbers of offenders caught in the act. Despite the denials of PAVED's designers, this remains the benchmark of efficiency for police in an era of target-based objectives policies.[107]

For the opponents of such systems, this information may seem reassuring. But in reality, even if the 'predictive policing' trend appears to have lost some momentum in France, research into and development of automated decision-making systems for the police continues. Substantial sums of money continue to be spent to meet the stated ambition of "taking the Ministry of the Interior to the technological frontier," as expressed in the 2020 Internal Security White Paper.[108]

Given the priority ascribed to techno-security approaches, PAVED could be reactivated or replaced by other systems in the near future. As for Edicia, in recent months the company has been considering incorporating new sources of data from social media into its 'predictive' module. This is what was originally planned by the designers of M-Pulse. 'Predictive' policing therefore remains a current issue.

## 3.7 Data collection and use: serious shortcomings

The collection and use of sensitive and identifying information in these systems is also a serious issue. For example, the 'predictive' module proposed by Edicia for Smart Police feeds on categories of geolocated data that are gradually generated as the software is used: complaints, incident reports, unofficial information and even rumours.

107. Cécile Godé, Sébastien Brion, and Amélie Bohas, "The Affordance-Actualization process in a Predictive Policing Context: insights from the French Military Police," in European Conference on Information Systems (ECIS) (Marrakech, Morocco, 2020). Available at: *https://hal.archives-ouvertes.fr/hal-02500125*

108. The White Paper proposed devoting 1% of GDP to internal security missions by 2030, representing an expected increase of around 30% in the Ministry's budget over the decade. Ministry of the Interior, "Livre blanc de la sécurité intérieure" (Paris: French Government, November 16, 2020). Available at: *https://www.interieur.gouv.fr/Actualites/L-actu-du-Ministere/Livre-blanc-de-la-securite-interieure*

Edicia, like other developers of 'predictive' systems, claims that no personal data is processed. Presumably, the input data – insofar as the datasets in question are indeed likely to contain personal data – is redacted or at least anonymised before being fed into these systems. The documents we consulted concerning Edicia's software do mention anonymisation procedures. However, the interface of the 'Field' and 'Active Vigilance' tabs of the Smart Police software seem to give officers in the field free rein to record any information they deem useful and relevant. The software also enables ID cards to be scanned.

Given the sensitivity of this data, even if it is not actually processed as part of the 'predictive' module, the system's compliance with data protection rules should be carefully assessed. However, we are not aware of any legal analysis carried out by Edicia or its clients regarding the risk of illegal profiling. When questioned on this point via a CADA request in March 2022 and again in November 2023, the CNIL told us that it had never received or produced any document relating to Edicia's Smart Police software.[109] This indicates that the data protection authority has never taken an interest in it. This in itself raises concern, considering the software is used by several hundred municipal police forces across the country.

Finally, although administrative police powers for areas deemed 'at risk' by 'predictive' systems can be considered 'individual administrative decisions', the requirements set out in French jurisprudence on algorithms should be respected.[110] These requirements include the possibility of administrative appeal for impacted people.

In addition, there are transparency obligations imposed by law, notably the 2016 'Digital Republic' law.[111] These legislative and jurisprudential requirements do not appear to be met when it comes to 'predictive' policing systems. Firstly, there is there no significant, proactive attempt to inform citizens and other stakeholders about how exactly these systems work, apart from the occasional bit of information disseminated opportunistically when they are first deployed. Secondly, our use of the right to administrative information had extremely limited results. For the most part, we did not receive responses to our requests, particularly those sent to the Ministry of the Interior.

109. See the CNIL's December 2023 response: https://madada.fr/demande/demande_cada_relative_au_logicie#incoming-7651.
110. See the decision on the transposition of the GDPR (decision no. 2018-765 DC of June 12, 2018) and that on Parcoursup (decision no. 2020-834 QPC of April 3, 2020).
111. On the legal obligation for transparency around public algorithms, see: Loup Cellard, "Les demandes citoyennes de transparence au sujet des algorithmes publics," Research Note (Paris: Mission Etalab, July 1, 2019). Available at:https://www.loupcellard.com/wp-content/uploads/2019/07/cellard_note_algo_public.pdf.

Institutional control of these policing systems therefore appears to be totally lacking. In this respect, the 'predictive' policing systems analysed in this report are not an exception. Instead, they provide yet another example of a recurring theme in the history of state surveillance: the chronic inability of the rule of law to provide an effective framework for the police's unquenchable thirst for new surveillance technologies.

# Conclusion: prohibit 'predictive' policing

For many years now, 'predictive' policing systems in France have been tested, and even permanently deployed, in almost total opacity. Most of them use biased police or other data, are based on dangerous criminological theories, and threaten civil liberties and human rights. The information gathered in this report shows that, in their very principle, most of these automated decision-making systems reinforce police surveillance and repression of groups who are already impacted by structural discrimination. Under the guise of 'algorithmic objectivity,' these systems appear to allow police forces to abuse their powers.

The major legal principles governing state surveillance at the European level were reiterated in February 2023 by the German Constitutional Court, in a decision concerning the use of Palantir by the police of the Hesse state in Germany. The court emphasised in particular the requirement for a clear and accessible legal basis incorporating certain safeguards.[112] In France, these principles appear to be totally ignored in relation to the use of data-driven 'predictive' systems. The scarcity of information available prevents people from exercising their right to an effective remedy against these systems used by the police.

Similarly, the principles of non-discrimination and evaluation of 'predictive' policing algorithms to reduce the risk of bias – principles highlighted by the European Agency for Fundamental Rights – do not seem to be respected.[113]

112. In its decision, the German Constitutional Court draws a distinction between the initial collection of data supplied to the software and the subsequent algorithmic processing based on this data. In the Court's view, the latter results in the creation of new information about individuals, based on interconnections and cross-references that could not have been deduced simply from the initial collection. According to the Court, the creation of this new, more complex information generates a new and potentially more intrusive interference with rights and freedoms. Bundesverfassungsgericht, February 16, 2023, Automatisierte Datenanalyse, no. 1 BvR 1547/19 and 1 BvR 2634/20. Press release available at: https://www.bundesverfassungsgericht.de/SharedDocs/Pressemitteilungen/EN/2023/bvg23-018.html. Full decision available at: https://www.bundesverfassungsgericht.de/SharedDocs/Downloads/DE/2023/02/rs20230216_1bvr154719.pdf?__blob=publicationFile&v=1.
113. European Union Agency for Fundamental Rights, Bias in algorithms: Artificial Intelligence and Discrimination, 2022. Available at: *https://fra.europa.eu/sites/default/files/fra_uploads/fra-2022-bias-in-algorithms_en.pdf*

Proponents of these technologies continue to push the belief that AI or automated decision-making systems can make police more “efficient”, despite a blatant lack of evaluation, failure to comply with the law, and poor operational results. From our point of view, however, these systems merely automate injustice and police violence.

Furthermore, even though substantial financial investments are still being made, there is no real control, transparency or public debate surrounding these systems. While some of them have been abandoned or only minimally used by police forces, others continue to be developed, tested and integrated into operational practices.

It is urgent to prohibit their use, and carry out rigorous evaluation of their implementation, impacts and risks. Given the knowledge we have obtained, albeit partial, we believe that greater transparency will further demonstrate the futility and risks around these systems, and emphasise the need for a prohibition.

To mitigate the opacity that is deliberately maintained by the designers of these systems and the public authorities who use them, we invite you to submit any documents or materials that could provide a better understanding of how they work via our anonymous document submission platform.

An explanatory tutorial and a link to the platform are available at : *https://technopolice.fr/leak/*.
You can also send documents by regular mail to the following address: 115 rue de Ménilmontant, 75020 Paris.

If you find any factual or analytical errors in this report, please let us know by writing to contact@technopolice.fr. And to support this type of research in the future, you can make a donation to La Quadrature du Net: https://www.laquadrature.net/donner/.

# Bibliography


Agence API. “Le modèle éco d’Edicia bascule vers le contrat Saas”, October 30, 2019. https://agence-api.ouest-france.fr/article/le-modele-eco-dedicia-bascule-vers-le-contrat-saas.

Al Hassani, Jamal. “A Marseille, une carte dopée aux données pour gérer l’espace public”, *Journal du Net*, August 29, 2019. https://www.journaldunet.com/economie/services/1443369-marseille-carte-plateforme-big-data/.

“Arrêté du 14 avril 2009 autorisant la mise en œuvre de traitements automatisés dans les communes ayant pour objet la recherche et la constatation des infractions pénales par leurs fonctionnaires et agents habilités”. Accessed December 9, 2023. https://www.legifrance.gouv.fr/loda/id/JORFTEXT000020692173.

“Bias in Algorithms – Artificial Intelligence and Discrimination”, Fundamental Rights Agency of the European Union, December 2022. https://fra.europa.eu/sites/default/files/fra_uploads/fra-2022-bias-in-algorithms_en.pdf.

Benbouzid, Bilel. “La prévention situationnelle : genèse et développement d’une science pratique”. Doctoral thesis, Lyon 2, 2011.

Benbouzid, Bilel. “Quand prédire, c’est gérer : la police prédictive aux États-Unis”. *Réseaux* No. 211, No. 5 (November 16, 2018): 22156.

Biderman, Albert D., and Albert J. Reiss. “On Exploring the ‘Dark Figure’ of Crime”. *The Annals of the American Academy of Political and Social Science* 374, No. 1 (November 1967): 115.

Bonelli, Laurent. “Les modernisations contradictoires de la *Police Nationale*”. In *L’État démantelé*, 10217. Cahiers libres. Paris: La Découverte, 2010. https://doi.org/10.3917/dec.bonel.2010.01.0102.

Bourget, Annie and Kante Thierno. *Procédé et système de surveillance et de prévention de dysfonctionnement en sécurité territoriale*. Institut national de la propriété intellectuelle 3079952. Courbevoie, n. d. https://cloud.laquadrature.net/s/kTxZtfxqmyGYP32.

Cellard, Loup. “Les demandes citoyennes de transparence au sujet des algorithmes publics”. Research notes. Paris : Mission Etalab, July 1, 2019. http://www.loupcellard.com/wp-content/uploads/2019/07/cellard_note_algo_public.pdf.

Coquille, David. “Marseille s’offre ‘Oracle’ pour des présages sécuritaires”. *La Marseillaise*, June 9, 2017. http://www.lamarseillaise.fr/marseille/societe/61034-marseille-s-offre-oracle-pour-des-presages-securitaires.

*Courrier Picard*. “Sous le capot de la police predictive”, April 2, 2018. https://web.archive.org/web/20220301133732/https://www.courrier-picard.fr/art/88625/article/2018-02-04/sous-le-capot-de-la-police-predictive.

Dechaux, Delphine. “Cet incroyable éditeur nantais de logiciels qui améliore la sécurité des villes en numérisant la *Police Municipale* et la RATP”. *Challenges*, June 1, 2018. https://www.challenges.fr/high-tech/cet-incroyable-editeur-nantais-de-logiciels-qui-ameliore-la-securite-des-villes-en-numerisant-la-police-municipale-et-la-ratp_590969.

Didier, Emmanuel. “‘Compstat’ à Paris : initiative et mise en responsabilité policière”. *Champ pénal/Penal field*, No. Vol. VIII (June 11, 2011). https://doi.org/10.4000/champpenal.7971.

Dole, Marie-Sylviane. “Demande d’accès aux documents publics relatifs à la surveillance algorithmique de l’espace public marseillais”. Deputy Head of Legal Affairs, City of Marseille, July 6, 2020. https://data.technopolice.fr/fr/entity/yjhsdstl1a?page=1.

Epp, Charles R., Steven Maynard-Moody and Donald Haider-Markel. *Pulled over: How police stops define race and citizenship*. University of Chicago Press, 2014.

European Public Accountability Mechanisms. “France Public Accountability Index”, 2022. https://europam.eu/?module=country-profile&country=France#info_FOI.

Friendly, Michael. “The life and works of André-Michel Guerry, revisited”. *Sociological Spectrum* 42, No. 46 (November 2, 2022): 23359. https://doi.org/10.1080/02732173.2022.2078450.

Friendly, Michael and Nicolas de Sainte Agathe. “André-Michel Guerry’s Ordonnateur Statistique: The First Statistical Calculator?” *The American Statistician* 66, No. 3 (August 1, 2012): 195200. https://doi.org/10.1080/00031305.2012.714716.

Gauthier, Florian. “Prédire les vols de voitures ?” *Etalab* (blog), January 12, 2018. https://www.etalab.gouv.fr/predire-les-vols-de-voitures.

Germes, Melina. “Cartographies policières : la dimension vernaculaire du contrôle territorial”. *EchoGéo*, No. 28 (July 8, 2014). https://doi.org/10.4000/echogeo.13856.

Giménez-Santana, Alejandro, Leslie W. Kennedy and Joel M. Caplan. “Risk terrain modeling and the study of the physical determinants of criminal behavior”. *Cahiers de la sécurité et de la justice*, No. 47 (March 2019): 12634.

Global Right to Information Rating. “France Country Details”, September 2011. https://www.rti-rating.org/country-detail/.

Godé, Cécile, Sébastien Brion and Amélie Bohas. “The Affordance-Actualization process in a Predictive Policing Context: insights from the French Military Police”. In *European Conference on Information Systems (ECIS)*. Marrakech, Morocco, 2020. https://hal.archives-ouvertes.fr/hal-02500125.

Gosselin, Camille. “La police prédictive : enjeux soulevés par l’usage des algorithmes prédictifs en matière de sécurité publique”. Paris: IAU Île-de-France, 2019. https://www.institutparisregion.fr/fileadmin/NewEtudes/Etude_1797/Etude_Police_Predictive_V5.pdf.

Gros, Maryse. “Edicia associe big data et sécurité urbaine”. *Le Monde Informatique*, February 19, 2014. https://www.lemondeinformatique.fr/actualites/lire-edicia-associe-big-data-et-securite-urbaine-56624.html.

Harcourt, Bernard E. *Illusion of Order: The False Promise of Broken Windows Policing*. Harvard University Press, 2005.

Kante, Thierno and Philippe Leray. “A Probabilistic Relational Model Approach for Fault Tree Modeling”. In *Advances in Artificial Intelligence: From Theory to Practice*, ed. by Salem Benferhat, Karim Tabia and Moonis Ali, 15462. Lecture Notes in Computer Science. Cham: Springer International Publishing, 2017. https://doi.org/10.1007/978-3-319-60045-1_18.

Kitchen, Peter Francis. “Exploring the link between crime and socio-economic status in Ottawa and Saskatoon: A small-area geographical analysis”. Department of Justice Canada, Research and Statistics Division, 2006. https://link.springer.com/article/10.1007/s11205-009-9449-2.

La Quadrature du Net. “Le Printemps Marseillais dans la Technopolice ? Au fond à droite”, June 8, 2021. https://www.laquadrature.

net/2021/06/08/le-printemps-marseillais-dans-la-technopolice-au-fond-a-droite/.
La Quadrature du Net. “Notation des allocataires : l’indécence des pratiques de la CAF désormais indéniable”, November 27, 2023. https://www.laquadrature.net/2023/11/27/notation-des-allocataires-lindecence-des-pratiques-de-la-caf-desormais-indeniable/.
La Quadrature du Net – Technopolice. “La police prédictive progresse en France. Exigeons son interdiction !” *Technopolice* (blog), July 23, 2020. https://technopolice.fr/blog/la-police-predictive-progresse-en-france-exigeons-son-interdiction/.
Lecorps, Yann and Gaspard Tissandier. “PAVED with good intentions: an evaluation of the *Gendarmerie* predictive policing system”. Centre d’Économie de la Sorbonne (CES), Université Paris 1 Panthéon Sorbonne. Paris, September 2022.
Mehrotra, Dhruv. “The Maker of ShotSpotter Is Buying the World’s Most Infamous Predictive Policing Tech”. *Wired*. Accessed January 10, 2024. https://www.wired.com/story/soundthinking-geolitica-acquisition-predictive-policing/.
Ministry of the Interior. “Livre blanc de la sécurité intérieure”. Paris: French government, November 16, 2020. https://www.interieur.gouv.fr/Actualites/L-actu-du-Ministere/Livre-blanc-de-la-securite-interieure.
Ministry of the Interior and City of Marseille. “Convention de mise à disposition de données pour la plateforme Big Data de la Tranquillité Publique”, January 13, 2020. https://data.technopolice.fr/fr/entity/eaxer6u3c5?page=2.
Perrot, Patrick. “L’analyse du risque criminel : l’émergence d’une nouvelle approche”. *Revue de l’Électricité et de l’Électronique* REE 2014-5 SEE (December 1, 2014). https://www.researchgate.net/publication/274071556_L’analyse_du_risque_criminel_l’emergence_d’une_nouvelle_approche.
———. “Le renseignement criminel : de nouvelles perspectives contre la criminalité organisée”. *Revue Défense Nationale* 779, No. 4 (2015): 1115. https://doi.org/10.3917/rdna.779.0011.
———. “What about AI in Criminal Intelligence? From Predictive Policing to AI Perspectives”. *European Law Enforcement Research Bulletin*, No. 16 (August 14, 2017): 6575.
Perrot, Patrick, Valescant, Nicolas and Camara, Daniel. “Forecasting criminal patterns for decision making”. *URSI France* Journées scientifiques 2017 (February 1, 2017). https://www.ursi-france.org/fileadmin/journees_scient/docs_journees_2017/data/articles/000054.pdf.
Ridgeway, Greg. “Experiments in Criminology: Improving Our Understanding of Crime and the Criminal Justice System”. *Annual Review of Statistics and Its Application* 6, No. 1 (2019): 3761. https://doi.org/10.1146/annurev-statistics-030718-105057.
Sardier, Thibault. “Cartographie criminelle : surveiller et prédire”. *Le Monde.fr*, January 5, 2018. https://www.lemonde.fr/idees/article/2018/01/05/cartographie-criminelle-surveiller-et-predire_5237723_3232.html.
South, Scott J. and Steven F. Messner. “Crime and Demography: Multiple Linkages, Reciprocal Relations”. *Annual Review of Sociology* 26 (2000): 83106.
“Témoignage Ville de Marseille : Projet Big Data de la Tranquillité Publique avec Oracle Big Data”, 2017. https://www.youtube.com/watch?v=qYmbfQ-8BB4.
City of Marseille. “Création d’un outil Big Data de la Tranquillité Publique et

prestations d'accompagnement – Cahier des Clauses Techniques Particulières (CCTP)", January 2017. https://data.technopolice.fr/fr/entity/69yuoeous9u?page=24.
———. "Dossier de présentation – Projet Big Data de la Tranquillité Publique", October 7, 2019. https://data.technopolice.fr/fr/entity/b6mmkge146?page=1.
———. "État des lieux du projet d'Observatoire Big Data de la Tranquillité Publique", October 3, 2019. https://data.technopolice.fr/fr/entity/litu04kbx4e?page=2.

This report was authored by Eda Nano and Félix Tréguer, edited by La Quadrature du Net, Griff Ferris and Sofia Lyall. Graphic design and formatting by Marne.


The publication of this report was supported by Statewatch as part of a European project funded by the European & Society AI Fund.