



**HAL**  
open science

## Notes on building an open data portal for a French police force

Antoine Mazières, Alexandre Thomas, Romuald de la Cruz

► **To cite this version:**

Antoine Mazières, Alexandre Thomas, Romuald de la Cruz. Notes on building an open data portal for a French police force. 2022. hal-03705017

**HAL Id: hal-03705017**

**<https://hal.science/hal-03705017v1>**

Submitted on 28 Jul 2022

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Notes on building an open data portal for a French police force

Antoine Mazières<sup>\*1,2</sup>, Alexandre Thomas<sup>†1,2</sup>, and Romuald De la Cruz<sup>‡1,2</sup>

<sup>1</sup>*Gendarmerie Nationale, General Directorate (DGGN), Issy-les-Moulineaux, France.*

<sup>2</sup>*Public Interest Entrepreneurs program (EIG), Interministerial Directorate of Digital Affairs (DINUM), Paris, France.*

June, 2022

GENDARMERIE.INTERIEUR.GOUV.FR/BAROMETRE-NUMERIQUE

## Abstract

*Le Baromètre Numérique* is a dedicated open data portal developed by the French *Gendarmerie Nationale*. It provides data with respect to digital services maintained by or related to this civil service. These notes document some background elements and development directions of this project. We aim to update these notes as the project evolve.

**Keywords:** open data; open government; police; gendarmerie; France

## Context

### Open government

By the end of the 1990's, the significance of the Internet as a governance platform started to be widely acknowledged, and considered through various analogies with emerging open source and open data communities. Several public initiatives throughout the 2000's led to the *Open Government Partnership*<sup>1</sup> in 2011<sup>2</sup> which brought together 77 countries around principles and commitments towards digital transformation, bearing on a number of areas such as governance, justice, citizenship, health and the fight against corruption. Broadly speaking, open government endeavors aim at adapting democratic principles of public accountability and transparency into the digital age, and empowering citizens participation and contribution to the public decision making

\*antoine.mazieres@gmail.com

†hi@alxthm.com

‡romuald.de-la-cruz@gendarmerie.interieur.gouv.fr

<sup>1</sup><https://www.opengovpartnership.org/>

<sup>2</sup>Joined by France in 2014: <https://www.etalab.gouv.fr/gouvernement-ouvert/>

process.

Though generally supported by government doctrine and public opinion, these axioms still have to overcome numerous cultural and practical challenges to be implemented within the structures and processes of every administration and political entity.

### Public tech fellowships

One of the many pitfalls to navigate is the tight tech labour market where big tech and startup companies seem to have a greater advantage in alluring the relevant workforce, such as programmers, designers, data scientists and jurists. To get around the lack of manpower, a number of so-called *public tech fellowships* have been set up to attract necessary talents. Pioneered in 2012 by the *Presidential Innovation Fellows* program<sup>3</sup> in the United States, it has been followed by likely minded initiatives such as the *Public Interest Entrepreneurs* program<sup>4</sup> in France (2017, *Entrepreneurs d'Intérêt Général*, EIG), *Tech4Germany*<sup>5</sup> in Germany (2018) and the *Innovation Fellowships*<sup>6</sup> in the United Kingdom (2020)<sup>7</sup>.

These institutional set-ups are generally open through a competitive selection for well-funded short-term contracts with the mandate to swiftly explore and prototype high-impact services with a strong use of digital tools and methods. Participants usually have a high-level reporting line which contribute to the program prestige while helping to ease

<sup>3</sup><https://presidentialinnovationfellows.gov/>

<sup>4</sup><https://eig.etalab.gouv.fr/>

<sup>5</sup><https://tech.4germany.org/>

<sup>6</sup><https://no10innovationfellows.campaign.gov.uk/>

<sup>7</sup>The similarities between these various programs have been discussed by several members of their management teams during an international panel at the EIG 4<sup>th</sup> cohort *demo day*: <https://www.dailymotion.com/video/x82ilbt>

the relation with the administrations involved in the project.

The open data portal described here was developed in the context of the 5<sup>th</sup> cohort of the *Entrepreneur d'Intérêt Général* program, between September and June 2022, at the *Gendarmerie Nationale*.

## Gendarmerie Nationale

*Gendarmerie Nationale* (GN) is, along with *Police Nationale* (PN), one of the two main police forces in France. While PN has authority mainly in big cities, GN is the default police force in most of the rural and suburban areas which are less populated. As a result, GN's authority reaches, like PN, around 50% of the French population but covers 95% of the French territory.

The Gendarmerie inherits from policing institutions as far as from the middle-age (*Maréchaussée*), making it one of the most ancient French institutions. It has approximately 100,000 agents, mostly under military status, while being, since 2009, under the authority of the Minister of Home Affairs (*Ministère de l'Intérieur*).

With the digitalisation of their activities since the early 2000's, GN has gained a certain visibility in several open source communities by implementing on an institution-wide scale tools such as *Open Office* and *Linux*<sup>8</sup> and making contributions to several codebases<sup>9</sup>.

## Project

### From data intelligence to data transparency

Over the course of the last decade, GN has developed a number of digital services to interact with French citizens, such as a fully featured website, recruitment processes, a 24/7 online chat and numerous interfaces to fill administrative forms related to frauds or offences. The underlying systems have been developed along the way by different teams relying on a diverse set of technologies, sometimes with the support of external companies and third-party services.

The initial project<sup>10</sup> joined the EIG program with the perspective of developing a data intelligence platform for internal use. It entailed gathering and

<sup>8</sup><https://ubuntu.com/blog/la-gendarmerie-nationale-upgrades-85000-pcs-to-ubuntu-desktop-edition>

<sup>9</sup>Among which *LemonLDAP* and *OCS Inventory*.

<sup>10</sup><https://eig.etalab.gouv.fr/defis/cyberimp-ct/>

analyzing data from multiple sources in order to identify and represent relevant indicators for the national and territorial commands to monitor the above-mentioned digital services.

For a couple of months, we met with all the teams involved in the 15 data sources considered by the project. This *data diplomacy* helped us to lay down the information available and the technical challenges to process it. What appeared is that, first, only aggregated data would be available in the short-term, yielding predefined variables that would make it difficult to craft indicators tailored to the command expectations. Second, and more importantly, that the data protection regulation implies a formal approval from the *National Commission on Informatics and Liberty* (CNIL) in order to centralize several data sources for a new purpose. This approval takes time and couldn't be sought in the short-term.

However, we realized that, given that we trim the available data from personal information, the legal hassle to process it for internal use would be equivalent to the one of releasing the data publicly. Therefore, instead of crafting an internal platform with limited value for the interested parties, we could endeavor developing a dedicated open data portal for GN.

## Open data & police

As it joined the *Open Government Partnership* in 2014, the French government released two consecutive *National Action Plans* (2015-2017, 2018-2020)<sup>11</sup> to lay down the principles on which it would foster a "transparent and collaborative government". In 2016, the *Law for a Digital Republic*<sup>12</sup> enforced some of these commitments, including making it mandatory for administrations to publish their data.

One caveat to this obligation is to protect personal information and the proper functioning of the administration. Police forces are both processing sensitive personal information and their activity require a certain amount of discretion if not secrecy, usually enforced by law. Therefore, police forces have generally initiated their journey to opening their data with caution, and there are not as many of them throughout the world who have opened their data as other types of civil services may have.

<sup>11</sup><https://www.etalab.gouv.fr/plan-daction-national/>

<sup>12</sup>*Loi n°2016-1321 pour une République numérique.* <https://www.legifrance.gouv.fr/dossierlegislatif/JORFDOLE000031589829/>

However, numerous initiatives are being carried out, such as the *Police Data Initiative*<sup>13</sup> in the USA which gathers the publication of 120 police forces – more than 200 datasets – and puts forward the open-ended “purposes of enhancing trust, understanding, innovation, and the co-production of public safety”. In the UK, a national dedicated platform<sup>14</sup> offers both datasets and APIs which enable users to access programmatically the latest information about a range of topics spanning from street-level crimes to neighborhood policing priorities. Other salient instances include the platforms from Toronto<sup>15</sup> and New Zealand<sup>16</sup> police forces.

In France, since 2019, administrations must publish their data on [data.gouv.fr](https://data.gouv.fr) which – as of today – indexes more than 40,000 datasets, several hundreds of which are related to police forces and published by the statistical service of the Minister of home affairs<sup>17</sup>.

### Target audiences & interfaces

We developed the *Baromètre Numérique* bearing in mind a number of possible audiences and use cases, among which:

- Citizens willing to know more about the underlying activity of the digital services they are using to interact with an administration.
- Journalists and researchers willing to report on these digital services or the underlying facts their activity are a proxy to.
- Other administrations or companies willing to embed this data into their own services.
- GN’s agents and services themselves as a way to evaluate and emphasize their work.

Meeting, at least partially, the requirements of these different use cases requires to avoid some common options considered for data portals, for instance:

- A list of datasets available for download along with a short description which tends to require that the user will have the time and knowledge to exploit the data with some specialized software.

<sup>13</sup><https://www.policedatainitiative.org/>

<sup>14</sup><https://data.police.uk/>

<sup>15</sup><https://data.torontopolice.on.ca/>

<sup>16</sup><https://policedata.nz/>

<sup>17</sup>SSMSI, <https://www.interieur.gouv.fr/Interstats/Actualites>

- A complex dashboard with numerous and complex visualizations which might necessitate a certain amount of expertise to interpret the data or even to identify the topic at hand.

In order to try meet various degrees of police and tech savviness, we decided to provide both a minimalist visual interface and a programming interface (API).

Each page deals with only one digital service and is composed of simple isolated blocks, starting with a highlighted key figure and description. The whole site is used composed of only 3 types of visualisation, an interactive map of France and 2 bar charts, one for categories and one for timelines. A simple tag system enables the user to modify the underlying query for a set of different periods (week, month, year). A simple calendar to manage this input should enable finer queries from the web interface.

The API enables advanced usages. Similarly to a static file or a database, it allows users to access all the data but adds also the possibility to programmatically access and embed the specific up-to-date elements from a third-party program or website. However, this interface doesn’t combine well with the use of traditional spreadsheet softwares and we will automate promptly regular dumps of the data on the platform [data.gouv.fr](https://data.gouv.fr).