The State of Open Data

Walker, B., Davies, Tim

Published by African Minds

Walker, B. and Tim Davies.
The State of Open Data: Histories and Horizons.
Project MUSE. muse.jhu.edu/book/67112.

For additional information about this book
https://muse.jhu.edu/book/67112

For content related to this chapter
https://muse.jhu.edu/related_content?type=book&id=2375471
Key points

- Some of the earliest open data experiments revolved around crime data and were driven by public and journalistic interest in local crime data; however, the open data community related to crime and justice data remains one of the least developed.

- Open data work in the crime and justice domain faces particular challenges related to privacy, legacy systems, and interoperability, and often involves working with some of the most conservative institutions.

- Donors and international organisations have increasingly recognised the potential links between open data and the crime and justice sector, but there are many cultural and coordination barriers to be overcome.

- A strong and sustainable judicial open data ecosystem has the potential to create more transparent and accountable judicial institutions and to improve the quality and effectiveness of judicial public policy, leading to greater access to justice and safer environments for all.

Introduction

In May 2005, a month before the official Google Maps application programming interface (API) was launched, ChicagoCrime.org was launched, a pioneering experiment that took crime data from the Chicago Police Department and presented it on an interactive map. Not only did this inspire a plethora of diverse mapping mash-ups,1 but it also sparked many other data-driven crime maps and acted as a key reference point for early open data arguments.2 Yet despite this strong beginning, crime and justice probably remains one of the least developed sectors for open data. Outside of the publication of administrative data by police forces, usually at the local level to enable crime incident mapping, the release of open data from judiciaries or other entities
within the justice system remains rare, and governments are generally more reluctant to open up crime and justice data to the public. According to the last edition of the Open Data Barometer, only 17% of surveyed governments had, by 2017, made any crime data available to the public as open data.

There is, however, a growing awareness of “open justice” as a public good and the need to apply open data principles to enhance transparency, accountability, and citizen participation related to the activities of the government agencies dealing with crime and justice matters. Evidence of this movement toward “open justice” can be seen in the growing number of worldwide judicial commitments included by member countries of the Open Government Partnership (OGP) in their National Action Plans. In 2011, only two out of a total of 170 commitments delivered by member states related to the judiciary, while the last three years have seen an increase to 63 (16 in 2015, 25 in 2016, and 22 in 2017). Of the total of 100 justice-related commitments delivered within the OGP system since 2011, 24 are based on the use of open data (see box, Examples of OGP justice-related commitments based on the use of open data).

The importance of open crime and justice data derives from the need to reinforce transparency and accountability. Crime and justice institutions have historically been seen as rather aloof institutions, detached from social influence; however, these institutions take actions and make decisions that should really not be considered any differently from other public institutions with regard to the need to be transparent and subject to constant public scrutiny.

Government is often divided into three branches: legislative, executive, and judicial. Open government data programmes have predominantly focused on the executive, which has responsibility for the delivery of government services and the implementation of legislation, leading to the release of crime data and crime mapping from police institutions or ministries. However, there has been much less focus on the availability of open data from the judicial branch of government.

This must be addressed as a core issue of democracy as the justice system should be a citizen-centred public service, where decisions are actually made by civil servants who are entrusted with the task of observing the law but are in no way above it. Greater levels of open data from the judiciary can help the system become not only more transparent and accountable, but also more efficient. Open data should inform judicial public policy. While, at present, policies are often designed from the top down and can result in poor quality services, the use of open data to build sound judicial policies through data analysis and citizen engagement will provide more efficient judicial services. Jimenez-Gomez recently described the worldwide state of the art of “open justice” initiatives more broadly and identifies the use of open data as a core element that should be taken into account to enhance the accountability of the courts.

There are many elements that go into open justice, often involving the innovative use of technology in the crime and justice space. This chapter is intended to be an examination of the evolution of open data specifically pertaining to crime and justice, and is not intended to include the additional analysis of policies related to civic participation or technology other than those related to open data.
Examples of OGP justice-related commitments based on the use of open data

France (2015), National Action Plan 1, Commitment 12: Open Legal Resources
This commitment received a starred rating after evaluation by the OGP’s Independent Reporting Mechanism due to its potentially transformative impact. It includes further developing the provision of legal and legislative resources as reusable open data and deepening citizen participation in developing innovative services and open source tools to facilitate the understanding and preparation of legislative texts, as well as in the drafting (avant-projet de loi) of the Digital Bill.

Spain (2017), National Action Plan 3, Commitment 4.1: Open Justice in Spain
This current commitment focuses on advancing open data as an instrument for achieving openness in Spain's judicial branch. It seeks to promote the citizen's right to access judicial information, including the initial steps required to transform the existing model of judicial statistics into a new system based on open data with improved characteristics regarding the quality of data, its collection, and management.

Lessons learned: Open crime and justice data

The most common sources of open data in this domain are the government agencies responsible for delivering services and implementing policies related to crime and justice. Hence, it makes sense that a vast majority of open data initiatives are national projects driven by institutions of the executive branch and the judiciary, with a few of them carried out by international organisations such as the European Union (EU) or the United Nations Office for Drugs and Crime (UNODC). While there is a not negligible amount of crime and justice data collected by private sector institutions, such as law firms, the potential contribution of these sources is yet to be realised.

Bargh, Choenni, and Meijer have accurately identified “three typical challenges” for the implementation of open data in the judicial field, highlighting privacy, legacy, and interoperability as significant challenges that should be taken into account in further development.9

Privacy relates to the required balance between the transparency of data and privacy for the real-life persons whose sensitive attributes, such as names, birth dates, crime types, or judgments, must be protected by removal or anonymisation.

Legacy refers to the very nature of legal data and the semantic evolution over time caused by continuous changes to rules and regulations. New crimes under the law need to be codified and old crimes can have their names changed or be redefined. Managing legacy data also becomes a challenge when government reform initiatives involve switching to new IT landscapes, requiring the migration of large amounts of accumulated data (data historically stored on paper) and then transferred to newer electronic systems. In order to make this data open and reusable, the importance of effective independent management of legacy systems cannot be underestimated.
Finally, the challenge of **interoperability** alludes to the necessity of ensuring that different sets of data, gathered by a large number of different agencies, be collected, stored, and then released using standardised criteria and processes, allowing the data to be integrated and combined with data from external sources. The justice system also needs to advance the use of unique identifiers that would make it easier to connect data across institutions and avoid redundancies in data collection between multiple partners across the judiciary who may be recording the same information.

The structure of the judiciary within federal governments deserves a special mention as the existence of national and sub-national levels (involving different judicial systems within one country) requires complicated inter-institutional coordination, making synchronisation and interoperability especially difficult to accomplish. Collaboration between different branches of government also presents challenges to reforms driven by the openness agenda as the transversal interaction required is perceived, in some cases, as a threat to the separation of powers (i.e. judicial independence).

Political and cultural barriers are still common hurdles for the implementation of the open data agenda in the public sector. Those barriers tend to be even higher in the case of institutions dealing with crime and justice (law enforcement agencies, the judiciary), which are traditionally some of the most conservative and independent institutions. As Roberto Gargarella notes, this is based on a conception of impartiality that holds dear the idea of isolated reflection by an individual (or a small elite of individuals) as a requisite for making correct or unbiased decisions.10

Another challenge is created when police and justice institutions lag behind in terms of technological capacity (i.e. hardware, software, skills, and expertise). Universities and scientific agencies could, and should, play a key role in building capacity for the use of data and emerging technologies in the police and justice sector.

Additionally, the involvement of civil society is still very limited in this field. Few civil society networks have projects looking specifically at open data in the crime and justice domain, with even fewer being large enough to be known internationally. However, there are some emerging examples of civil society organisations (CSOs) working independently or in collaboration with government agencies, including Measures for Justice in the United States (US), OpenGiustizia in Italy, La Nación Data in Argentina, and the Justice Data Lab11 in the United Kingdom (UK).

**Measures for Justice**12 is a civil society initiative launched in 2011 that has developed a data-driven set of performance measures aimed at assessing and comparing different aspects of the criminal justice system in state jurisdictions of the US. The analysis, using data extracted from administrative case management systems, covers three main categories: fiscal responsibility, fair process, and public safety.

**OpenGiustizia**13 was a project focused on organisational innovation and optimisation for the Court and the Public Prosecutor of Napoli, developed by three Italian universities and financed by the EU’s Social Fund between 2007 and 2013. Among the project’s objectives was the creation of interoperability within the system’s databases and the provision of tools for accountability and performance evaluation.

**La Nación Data**14 is a data journalism initiative which has been underway since 2012 by one of the main newspapers of Argentina. It consists of a news portal and blog based on data collected...
from various sources. It makes an intensive use of open crime and justice data, delivering content on themes such as femicide, high-profile judicial cases, and the penitentiary system.

**Justice Data Lab** is a service run by the Ministry of Justice of the UK and New Philanthropy Capital. Set up in 2013, it is aimed at organisations providing offender rehabilitation services. It uses administrative data on re-offenders to conduct on-demand impact evaluations, so that these organisations can assess the actual impact of their work through data-based evidence.

For most of these initiatives, data availability and interoperability remain a challenge. Measures for Justice, for example, covers just six US states at present. The next section will explore the different kinds of data that could or should be available on crime and justice, and the various actors involved in creating and using it.

### Making crime and justice open data available

The primary focus of opening data within the crime and justice sector is on three main categories of information:

1. Case data: information on judgments and court rulings issued by crime and justice institutions (e.g. courts, tribunals, etc.).
2. Jurisdictional data: performance and activity data from crime and justice agencies, such as statistical data related to cases, reported crimes, arrests, citizen complaints, etc.
3. Structural data: information on the internal characteristics of crime and justice agencies, such as their organisation, their internal processes, how they allocate their budget, infrastructure, rules of procedure, staff and salaries, procurement, etc.

Data produced or collected by institutions within the crime and justice system is generally made available in three main ways:

1. As primary data (i.e. unprocessed, as it was collected at the source) in downloadable datasets or files (CSV, XML, DOC, XLS, PDF).
2. As aggregated statistical data (i.e. as processed, and anonymised if necessary, data) in the form of downloadable files or datasets.
3. In aggregate form, but as graphical presentations either in static visualisations of statistical data or through the use of user-facing data visualisation and analysis tools.

As the box below illustrates, while jurisdictional and structural data may originate with either the judicial or executive branches of government, case data tends to be solely within the province of the judiciary.
The open crime and justice open data ecosystem

Using a data ecosystem mapping methodology, this diagram represents the crime and justice open data ecosystem as described in this chapter. Unbroken lines represent constant interaction between the actors involved, while dotted arrows represent a direct but non-continuous bond. Information producers (e.g. the judiciary, ministries of justice, the police) are often also active consumers and users of the information produced. Stakeholders, such as academia, CSOs, and data journalists, act as intermediaries, using raw data and transforming it into user-friendly information products for a broader range of users (citizens).

The way in which open crime and justice data are combined and delivered to users can vary substantially. Table 1 gives 18 examples of data projects, although, for the sake of this chapter, detailed analysis is restricted to four illustrative cases: Openjustice (US), Data.police.uk (UK), Datos.jus.gob.ar (Argentina), and ECourts (India).

Openjustice is an open data project developed by the Office of the Attorney General of the Department of Justice of California (US) to establish a criminal justice data portal which was first launched in 2015. It currently delivers jurisdictional data from all enforcement agencies across the State of California, covering crime, deaths in custody, hate crimes, homicides, juvenile court and probation, citizen complaints, and the use of force. Structural data on the portal includes lists of law enforcement and criminal justice personnel, as well as county-level contextual data (educational attainment, income, poverty, and unemployment levels) for each county. The data is made available as downloadable datasets and through visualisation tools.

Data.police.uk is an open data portal maintained by the Home Office of the United Kingdom that provides data about crime and policing in England, Wales, and Northern Ireland.
Table 1: Open crime and justice data projects around the world

<table>
<thead>
<tr>
<th>INITIATIVES</th>
<th>TYPE OF DATA</th>
<th>FORMATS</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Cases</td>
<td>Jurisdictional data</td>
</tr>
<tr>
<td>OpenJustice (US)</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>data.police.uk (UK)</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>datos.jus.gob.ar (Argentina)</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>ECourts (India)</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Measures for Justice (US)</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Mapa del Delito CABA (Argentina)</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Datos Abiertos del Poder Judicial de Costa Rica (Argentina)</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Data Portal Singapore’s Public Data (Singapore)</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Productivity Commission (Australia)</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Dados Abiertos MPRS (Brazil)</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Judicial Department (Russia)</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Statistics Canada Crime and Justice (Canada)</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>The Judiciary (Liberia)</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>data.unodc.org (UNODC)</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>ISS Crime Hub (South Africa)</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Otvorené Südy (Slovakia)</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Eur-lex (EU)</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>De Rechstpraak (Netherlands)</td>
<td>X</td>
<td>X</td>
</tr>
</tbody>
</table>
It was launched in 2013\(^3\) and delivers jurisdictional data on reported crimes and all kinds of police activity, including drug seizures, the issuance of firearms certificates, breath tests, or the setting up of cordons under the Terrorism Act. It also contains structural data on the police workforce, procurement, salaries, etc. Both primary and aggregated statistical data is available, and the update frequency is either quarterly or annually depending on the subject matter.

Datos.jus.gob.ar\(^3\) is the open data portal of the Ministry of Justice and Human Rights of Argentina, containing overall data on the country’s justice sector. The portal was launched in 2016 and offers data on a range of jurisdictional activities, such as the delivery of pre-judicial mediation and the provision of access to justice, as well as information on criminal policy, the prison system, and structural data on institutions of the judicial branch and the Ministry of Justice. Primary and aggregate data are available as downloadable datasets as well as via visualisation tools. The update frequency of datasets depends on the subject matter, ranging from daily or monthly to annually.

ECourts\(^3\) is a service provided by the Ministry of Law and Justice and the Supreme Court of India. Online since 2013, it contains real-time judicial data for all jurisdictions subject to the Indian judiciary. It aims to serve as a dynamic source of information on the judicial system. It is based upon a “National Judicial Data Grid”, which works as a nationwide data warehouse for case data and aggregated data delivered through visualisations.

Additional examples of open crime and justice data initiatives are highlighted in Table 1, classified according to the categories mentioned previously.\(^3\)

**The potential impact on social and economic development**

Open crime and justice data is expected to play a key role in measuring and delivering progress in terms of social and economic development. Although the UNODC has been working on crime and justice statistics for many years, the United Nations 2030 Agenda for Sustainable Development places a fundamental importance on open data at all levels to promote accountability and inclusive decisions, to support reductions in crime and violence, and to improve access to justice for all over the next 11 years. Although data will play a vital role in showcasing national progress toward over 169 global targets encompassed within the Sustainable Development Goals (SDGs), it will also allow decision-makers in all three branches of government to be able to rely on quality information for the design of public policies, based on evidence, to achieve those global targets. Key SDGs for crime and justice data include SDG 16, aimed at the reduction of violence, the reduction of organised crime, the development of effective, accountable, and transparent institutions, and ensuring access to justice and public information, and SDG 5, which focuses on gender equality and the total elimination of violence against women and girls.

With regard to SDG 5, two specific examples of effective initiatives should be noted: the provision of primary open data on sexual offences by the Colombian government through their Open Data Portal\(^3\) and the specific section on gender issues of the Open Judicial Data Portal of the Ministry of Justice of Argentina, datos.jus.gob.ar,\(^3\) where primary data is made available on femicides, human trafficking, and assistance granted to victims of violence.
The potential of open crime and justice data with regard to the SDGs will probably have a crucial impact in the future allocation of resources and funding for associated projects and initiatives. International organisations, such as The Hague Institute for Innovation of Law or the Latin American Open Data Initiative (ILDA), are already orienting their funding priorities in this direction as are other significant actors like the Open Data Institute, Transparency International, and the Open Society Foundations. Crime and justice open data is also increasingly on the agenda for key international organisations that are pushing for open government-oriented reforms in the public sector, including the International Development Research Centre, the OGP, and mySociety, among others.

Conclusions

An enabling environment is currently emerging for the spread of open data by and between crime and justice institutions. International organisations and governments have begun to consider crime and justice data as a raw material to use in implementing and evaluating public policies. At the same time, data journalists, academia, and CSOs are learning how to use open data to promote more transparent and accountable judicial institutions.

There are still, however, many barriers to the implementation of quality open data initiatives and specifically to the extended use of judicial open data. The main barriers are traditional cultural and political forces against openness within the judicial system, the lack of adequate financial and human resources invested in capacity building, and current inadequate or restrictive legal frameworks, including those that create a barrier to publishing data from judicial cases. Another hurdle to overcome is the difficult but necessary coordination of the various public institutions producing judicial data, as well as the lack of consistent standards around the production and publishing of judicial information and the relatively weak expertise of civil society and other actors in analysing the data.

Ultimately, two main actions are required in order to strengthen the judicial open data ecosystem. First, at the institutional level, governments should include the judicial sector within their access to information laws and open data policies and regulations. At the same time, justice institutions should commit to country-wide and transversal open data strategies. These strategies must take into account open judicial data intermediaries, such as academia and CSOs. It is recommended that one judicial institution in the country (e.g. the Ministry of Justice, the Supreme Court, or the Judicial Council) takes on the leadership role and coordinates the development and implementation of open data policies and plans. Additionally, judiciaries should set goals, targets, and indicators for justice delivery, and the resulting performance data should be available and evaluated through open data.

Second, concerning the use of open judicial data, each judicial system must establish a balance between publication and privacy protection. Privacy should not be used as an excuse for avoiding openness. Governments and international organisations should promote the use of open judicial data through different tools (curricula, hackathons, journalism, etc.) and public participation mechanisms should be put in place to assess and set priorities for the data release process. We also recommend setting up open judicial data portals to contain the totality of available data for
each judicial system. Open judicial data should not only be provided in the format of open datasets, but also through visualisations and data stories in order to reach a wide variety of users. Advancing the interoperability of the data held in the systems of numerous institutions producing open judicial data is a must. Governments, working with leading international organisations, should promote the definition and adoption of specific standards for open judicial data. International organisations should also promote the creation of open judicial data networks and working groups, as well as the participation of open judicial data experts and leaders in relevant conferences and debates wherever they are taking place. With success in these endeavours, the next decade should see the ongoing development of a strong and sustainable judicial open data ecosystem that can enable more transparent and accountable judicial institutions while delivering more effective access to justice and safer environments for all.

Further reading


Open Data Sectors and Communities | Crime and Justice

About the author

Sandra Elena is an open data and open government expert based in Buenos Aires. She currently coordinates the Open Justice Program, Ministry of Justice and Human Rights, which is implementing the first open data initiative in Argentina’s judiciary. You can follow Sandra at https://www.twitter.com/sandra_elena1 and learn more about her work at https://datos.jus.gob.ar.

How to cite this chapter


This work is licensed under a Creative Commons Attribution 4.0 International (CC BY 4.0) licence. It was carried out with the aid of a grant from the International Development Research Centre, Ottawa, Canada.

Endnotes


7 https://www.opengovpartnership.org/starred-commitments/open-legal-resources


12 https://measuresforjustice.org/

13 http://www.opengiustizia.it/

14 https://www.lanacion.com.ar/data


16 http://www.opendataresearch.org/emergingimpacts/methods.html


18 https://openjustice.doj.ca.gov/

19 https://data.police.uk/

20 https://measuresforjustice.org/

21 https://mapa.seguridadcidad.gob.ar/

22 https://datosabiertospj.eastus.cloudapp.azure.com/


25 http://dados.mprs.mp.br/dados_abertos/

26 http://www.cdep.ru/


28 http://judiciary.gov.uk/

29 https://data.unodc.org/

30 https://issafrica.org/crimehub

31 https://otvorenesudy.sk/

32 https://eur-lex.europa.eu/

33 https://www.rechtspraak.nl/


35 http://datos.jus.gob.ar/

36 http://ecourts.gov.in/ecourts_home/

37 Information available at Open Data Portals as of 13 March 2018.

38 https://www.datos.gov.co/Seguridad-y-Defensa/Delitos-Sexuales-2016/3j7m-zgyi/data

39 http://datos.jus.gob.ar/pages/datos-de-justicia-con-perspectiva-de-genero