DATA, AI AND GOVERNANCE

Open Data: a Public Resource for Digital Transformation

Lourdes Muñoz



Araya Peralta

Open data represents an advancement for democracy, owing to the fact that, by making such data truly public, all citizens are afforded access to information that is both useful and usable. At the same time, open data serves as an excellent resource for the development of applications and services with a high level of social value.

The concept of open data refers to making data public, in such a way that they can be openly accessed and freely re-used. Open data is one of the three spheres of activity of open government.

The positive impact of open data on society can be summarized and divided into four pillars: the improvements undergone by governments themselves; its contribution to the development of improved solutions for public issues; the empowerment of citizens; and the creation of fairer economic opportunities.

Open-data strategies improve governments by providing useful information for the fight against corruption, by increasing transparency, and by improving public services and the distribution of resources.

Moreover, open data is playing an increasingly important role in solving major public issues. It provides us with a new approach to assessment while enabling a more direct and collaborative form of participation that is based on the data themselves.

> Open data represents an advancement for democracy, because, by making such data truly public, all citizens are afforded access to information that is both useful and usable

The empowerment and education of citizens is another consequence of making data more open. For this reason, it is important that access to the public data held by institutions should be considered a right.

Open data, accompanied by new forms of communication and access to information, enables decisions to be made by citizens who are not only more informed, but also better informed. It also promotes or allows for new forms of social mobilization.

The positive impact of open data on governance brings with it social and collective benefits which, by themselves, justify the public policies on open data that are awaiting implementation in our country. However, it is also important to demonstrate the economic potential of open data, because committing to this concept means promoting a fairer economy.

From an economic perspective, the use of open data provides the foundation for a large number of innovative activities and potential products and services, which will result in increased economic growth and more jobs. Making data open and available, as a public resource that can be accessed and re-used, is a strategy for generating economic opportunities.

Indeed, it is a public strategy that takes the “open” ethic as its starting point and aims to make the opportunities offered by the Internet become reality. The systematic publication of open data and the creation of spaces that offer structured information (and facilitate its transformation into knowledge) represents the public provision of the most valuable form of input within the context of a networked, information-based society.

Forward-looking legislation

The existing legislation is both forward-looking and demanding. The Catalan Transparency Act (Law 19/2014 of 29 December on transparency, access to public information and good governance) is an ambitious law, although it has not yet been fully implemented.

It specifically tackles open data and access to public information and incorporates the so-called “right to petition” as a subjective right. As such, it recognizes the right to request and obtain public information, otherwise known as “passive transparency”.

It also incorporates obligations related to “active transparency”, i.e. the publication of public information or the adoption of open data on the part of public administrations, in order to promote transparency and enable citizens to oversee government activity.

Specifically, the law stipulates that “public administrations must provide citizens with the information they require to understand and evaluate public management”, and establishes an obligation to publish data in a way that is both structured and systematic: “Transparency means playing an active role in communicating information. It does not simply mean publishing data: it means communicating information in a way that is structured and enables the management of said information”.

> Open-data strategies improve governments by providing useful information for the fight against corruption, by increasing transparency, and by improving public services and the distribution of resources

Of particular note is the Open Administration in Catalonia (AOC) initiative, which helps Catalonia’s municipalities to apply this legislation by creating a standardized solution in the form of a transparency portal. This portal includes a repository of global open data from a variety of supra-municipal sources. Currently, 100% of local authorities have a transparency portal, and 1,178 use the portal provided by AOC.

Moreover, a new platform has been developed that allows for the creation of an open-data portal to which administrations can upload their own data. The platform was developed with the active collaboration of a network of local councils, including Tarragona, Castellar del Vallès, Sant Just Desvern, El Prat de Llobregat and Sant Quirze del Vallès.

Nor should we overlook the pioneering initiatives that were implemented at the municipal level prior to the obligations imposed by the legislation. Gavà Municipal Council stands out in many rankings, owing to the fact that since 2014 it has provided an interactive serve that provides detailed information on an “invoice by invoice” basis, in a way that is both easy and comprehensible.

We are also in a good position within the context of Europe as a whole: in the European Data Portal’s Open Data Maturity Report 2019, which measures the development of open data, Spain occupies second place. Moreover, the new European directive will mark a new stage in the propagation of open data. More organizations will be obliged to implement the process and the data provided will be more relevant.

Those active in the sphere of economic re-use argue that it is necessary to move beyond the growth in the number of portals and datasets, and to start publishing data with values that allow said data to be re-used. This can be achieved by systematically applying the recommended parameters with regard to data quality (i.e. re-usability, accuracy, consistency, availability over time (both now and in the long term), completeness, use of standards, credibility of sources, precision, and the ability to be updated and understood). In the 2016 report by red.es characterizing the infomediary sector, the companies that make up the sector (and whose activities chiefly comprise the re-use of data) identified data

quality as one of their key needs. In the report titled The Re-use of Open Data in Spain (Abella et al., 2019), the authors concluded that the quality of the data had not improved in recent years.

Good practice facilitates the use of open data

A recent collaborative initiative has compiled and evaluated a series of projects (in Catalonia and worldwide) designed to facilitate the re-use of public open data (Open Data Ranking – Catalonia). The highest-ranked projects, which can serve as examples of good practice in facilitating the use and re-use of public information, include the following:

Search Engine for Official Documents and Information (CIDO). The CIDO service pertains to the Official Journal of Barcelona Provincial Council and aims to ensure access to public information while helping those public administrations that report to Barcelona Provincial Council comply with the requirement to provide active transparency.

Viladecans 360. Viladecans 360 is an online portal that provides access to all of the municipal information for the town of Viladecans, including downloadable files, maps and locations, data, and interactive statistics for Viladecans Municipal Council.

Crime map. This interactive map from the Ministry of the Interior provides information on crime in Catalonia and offers a visual overview of the day-to-day activities of the Mossos d'Esquadra, Catalonia's regional police force.

Budget implementation portal. This application, which has been developed by the Government of Catalonia and forms part of its General Intervention Service, aims to improve citizens' understanding of current progress regarding implementation of the budget.

Hermes infographics. This addition to the Hermes municipal statistics platform seeks to bring together as many local statistical indicators as possible on the same website, in order to support strategic and operational decision-making throughout the region. It forms part of the Government of Catalonia's Ministry of the Vice-Presidency and of the Economy and Finance.

The challenges facing Catalan institutions

The institutionalization of open data. In order to maintain an open-data initiative in a way that is sustainable, institutions must be equipped with a series of permanent tools. These could include: a regulation on open data that supports and legitimizes the opening-up of data; an open-data strategy or plan that clearly sets out its aims and a timetable for publication; and an integration plan for every organization that is a genuine source of a particular dataset, in order to enable said organizations to provide participative spaces for

data re-users and a technological platform that contains a catalogue of said data and allows it to be viewed.

Publication must be oriented towards re-use. Over 10 years ago, the leaders of the pioneering countries in this field (President Obama in the United States and Prime Minister Gordon Brown in the United Kingdom) launched their respective open-data portals within a few months of each other. Since then, we have learned that it is not enough to simply publish data: we need to actively promote the re-use of said data.

> Tools must be provided to transform the data into a social or economic asset or make them of value to citizens

The publication of data does not automatically transform those data into a social and economic asset. Tools must be provided so that individuals and organizations are able to transform the data into a social or economic asset or make them of value to citizens. If public data are not easy to process, then they are only a resource in theory and have no practical use.

Additionally, and more specifically, institutions must promote initiatives to guarantee access to and the use and transformation of public data. Otherwise, only those who are at the top of the digital transformation pyramid will be able to access them, and the data will no longer be a universal public resource.

When Prime Minister Gordon Brown unveiled the United Kingdom's open-data portal (data.gov.uk) in 2009, he explained that the aim of the initiative was to promote a fairer economy by providing access to the resource of public open data. "Data are the new raw material", he said, and open data is the most democratic kind. Indeed, the technological trends that have had the greatest impact on digital transformation are all based on data.

The United Kingdom, which is ranked number one in the global Data Barometer, has focused its openness strategy on the concept of data-based economic stimulation and on the use of data for research and innovation. To this end, it has created a specific body to fulfil this mission: the Open Data Institute.

The continued promotion of the re-use of open data has produced results. The 2015 report *Open Data means business: UK innovation across sectors and regions* identified 270 companies whose business models were based on open data. The companies were of varying sizes and operated in a range of different sectors: 54% were in the information and communication sector, 21% were in the field of science and technology, 11% provided services to business, while 5% were in the entertainment industry.

In Spain, once the data portals have been set up and are operational, open-data strategies must prioritize disclosure, education, and the creation of intermediary spaces that facilitate

the transformation of data into a social and economic asset.

Initiatives to promote the re-use of public data should incorporate a range of different action areas:

- Fostering the disclosure of information among SMEs and social organizations, in order to ensure fairness and promote an awareness of the subject.
- Organizing innovative events (e.g. data expeditions and datathons) aimed at finding solutions to social challenges through data. This can act as a strategy for promoting data technologies. Thus, open data becomes a vehicle for finding solutions: people, rather than open data, are the protagonists, and the focus is on the solutions that open data can provide.
- Promoting the use of data to generate open information and knowledge in a way that emphasizes disclosure. Using data visualization to explain social issues and challenges.
- Establishing active networks in order to generate an ecosystem that facilitates the capture of data for re-use, in order to promote innovation based on open data.

Institutional transformation in data-driven organizations

Data have become an infrastructure of public interest. They can be used to make decisions more quickly, to promote innovation, to improve public services and to empower people.

Through the use of Big Data, public administrations can improve their understanding of citizens' needs and predict their future requirements. In turn, this will enable them to manage their budgets more coherently.

Data-driven companies and organizations are based on information: data lie at the heart of their processes and the decisions they make. These organizations require absolute control over the information they process, which they achieve by merging and homogenizing the data they obtain from the different information sources they utilize. They also require intuitive and highly visual analytical methods in order to optimize and improve their decision-making processes.

The qualities that a data-driven organization needs to possess include the following:

- They must be underpinned by a strong data-analysis policy and have a data-oriented culture. They must define processes on the basis of indicators and have the support of a goal-oriented team.
- The data must be centralized, organized, continually updated and easily accessible.
- They must establish regulations that govern access to the data. Information must be governed in a sophisticated manner, ensuring that the "self-service" approach does

not come at the expense of security and that various forms of access are provided.
- They must have integrated tools for data analysis, and they must be able to integrate this analytical capacity into the platforms they traditionally and most frequently use.

Within this context, institutions are both producers of data (inasmuch as they publish open data) and consumers of data from other organizations, for the purposes of applying Big Data or artificial intelligence in order to generate data that enable smart governance. In order to meet the challenges faced by institutions with regard to the adoption of a data culture, said institutions will also require specific structures and professional profiles.

The Chief Data Officer (CDO) is the senior executive responsible for data processing and must lead the process of changing the organizational culture. The CDO must draw up and implement the organization's strategy with regard to data, standards, procedures and policies, and must manage teams of data experts at the corporate level.

In the United States, the Obama administration appointed the country's first Chief Data Scientist, DJ Patil, and the country's first national Chief Information Officers in 2009. The appointments were a landmark for innovation in the public sphere.

Closer to home, Barcelona City Council took a pioneering step forward in 2018 with the launch of the Municipal Data Office (OMD), which is responsible for the management, quality, regulation and utilization of the data owned or held by the City Council and all of its associated bodies. The OMD is charged with regulating and managing the city's data, and its aims include making municipal structures more dynamic and improving the efficiency and transparency of public policies with regard to data. Barcelona was the first city in Spain to create a body of this type.

The right to know – from a gender perspective

Without data equality, we cannot build an equal society. The majority of automated decision-making processes, in both the public and private spheres, are based on data processes and predictive AI-driven algorithms. For this reason, we must have access to data that provide a gender perspective, so that said data can serve as the basis for making better decisions: specifically, decisions that do not create additional gender gaps. We also need data on issues that specifically affect women, in order to deploy data technologies to help diagnose and tackle these issues.

Emily Courey Pryor, the Director of the Data2X initiative at the UN Foundation, has called for data with a gender perspective, in light of the fact that "without data equality, there is no gender equality". If specific data on the circumstances of women are not incorporated, it will be difficult to develop adequate strategies to achieve full equality for women. Indeed, the absence of such data could generate new gender gaps, forms of discrimination or biases (in the case of algorithms).

Applying a gender perspective to public data implies the provision of data that is broken down by sex, in order to compare information for men and women. It also implies the provision of data that refer exclusively to women, for reasons related to their gender (e.g. rates of cervical cancer), and data related to the roles that are assigned to each gender (e.g. data related to gender violence, prostitution or profession).

> Institutions must guarantee access to and the use and transformation of public data. Otherwise, the data will no longer be a universal public resource

To this end, CIBA, Santa Coloma Municipal Council and the Barcelona Open Data Initiative have launched an observatory for open data and gender, with a focus on sharing and calling for open data that includes women. Which data need to be gathered in order to tackle gender violence? Which indicators need to be broken down by gender in order to allow public open data to show the differences in the use of public services on the part of men and women? These are just some of the questions posed by this initiative.

If algorithms based on personal data can suggest specific products in accordance with our individual preferences, should we not focus our efforts on gathering data and creating algorithms that will help one half of our population attain equality with the other half? The systematic incorporation of a gender perspective in open data is of strategic importance if we are to meet the challenge of achieving real equality for women, given that public decisions are based, in large part, on such data. Moreover, the AI algorithms employed by private companies also make use of public data. Ultimately, public institutions must shoulder the significant responsibility of ensuring that women are not left out of the data.

The challenge posed by the new directive will impose more obligations on our institutions

A new European directive on open data and the re-use of public sector information has recently been approved. All EU Member States must transpose this directive into their domestic legislation before July 2021.

The directive contains new stipulations on the scope of open data, which will have a significant impact on the timetable for publication that must be implemented over the coming years. It also broadens the range of susceptible organizations to include publicly owned companies, provided these companies operate in certain areas of public sector contracting (e.g. water, energy, transport or postal services) or act as operators with an obligation to provide a public service.

Additionally, it broadens the range of data types that are considered to be of particular

relevance and must therefore be published and made available immediately (e.g. research data produced via publicly funded research, dynamic data and data produced in real time, the chief sources of which are sensors for smart-city projects). Furthermore, the scope for the publication of high-value data, which was initially limited to certain areas (e.g. geospace, the environment, meteorology, statistics, society and mobility), is extended to include those that are capable of generating socio-economic and environmental benefits or are capable of making innovative services more dynamic, and which must therefore be provided free of charge.



Lourdes Muñoz

Lourdes Muñoz és Enginyera Tècnica en informàtica i cofundadora de les iniciatives *Barcelona Open Data* i *Sheleader*. Ha tingut una àmplia carrera institucional, primer a l'Ajuntament de Barcelona (1999-2003), on va ser regidora de Polítiques de gènere, i després com a diputada al Congrés dels Diputats durant 11 anys (2002 a 2011 i 2014 - 2015). Com a parlamentària, ha format part de les comissions d'indústria, economia, innovació i igualtat. Com a portaveu del grup socialista sobre la societat de la informació, va participar en diverses iniciatives relacionades amb el dret de la societat a la informació, el dret en l'àmbit de les telecomunicacions, la reutilització d'informació pública i assessorament general dels col·legis professionals d'enginyeria informàtica i la promoció del govern obert i les dades obertes.