DATA, AI AND GOVERNANCE

The Open Data Strategy of the Government of Catalonia

Núria Espuny



Araya Peralta

Today, now just over two decades into the twenty-first century, the importance and potential of data is a well-acknowledged fact. Data, and the information that can be extracted from it, is a precious resource that can be used to generate social and economic value: data enables us to start new businesses, analyse patterns and trends, make data-driven decisions and solve complex problems. Public administrations, therefore, have a responsibility to ensure all citizens, all people and all businesses and organisations have the opportunity to harness this potential. It's what is known as open data.

The Directorate General for Transparency and Open Data at the Government of Catalonia is responsible for managing the process of opening public data and facilitating and encouraging its reuse, a task driven by our utmost belief in its benefits and potential, and carried out with the cooperation and support of all of the Government's ministries.

But what exactly does "opening data" or the concept of "open data" mean? In order to be categorised as "open", data must be easily accessible and reusable, at any time, free of charge, and in a format that allows it to be automatically processed by computers. The Government of Catalonia is a member of the Open Data Charter international organisation and, as such, has subscribed to the principles of its International Open Data Charter. Its mission is to instil the culture and practice of open data use in governments. The Charter says that public data must be:

- Open by default

- Timely and comprehensive
- Accessible and usable
- Comparable and interoperable

To make progress towards achieving these objectives, the Government has drafted its strategy for opening public data, which is regulated by Government Agreement 154/2018 and the 2019-2020 Open Government Plan. The strategy is based on three fundamental pillars:

- The opening of new data and, therefore, increasing the number of data sets available in the Government's catalogue of open data
- The improvement and promotion of data quality as an essential component of making it more useful
- Encouraging the reuse of data to return value to the citizenry

Let's discuss the first pillar: the opening of new data. One significant aspect of the Government of Catalonia's current open data strategy is the progress it represents with regard to the opening of new data. It establishes that, by default, all data that could be considered public must now be published in an open format. The move from publishing everything with a legal requirement to do so (*open by law*) to publishing everything that is not prohibited by law (*open by default*) is a considerable paradigm shift with significant implications for the organisation. These implications must be assumed and managed, not only at a budgetary level but also at a cultural and organisational level, which is often the most challenging part of any project. To change from a mindset that says *information is power* to aiming for *the socialisation of data*, is no small task. To support this pillar, one of the projects we're launching is the development of a single data inventory that keeps an orderly and descriptive record of each and every one of the data files available to the Government of Catalonia's Administration. The purpose of this measure is to identify any data that should be open and improve internal data management while avoiding inconsistencies and duplications. It also aims to facilitate interministerial and inter-agency data sharing.

> The move from publishing everything with a legal requirement to do so to publishing everything that is not prohibited by law is a considerable paradigm shift

The second pillar deals with another important aspect: managing the quality of data to ensure that it is reliable, accurate, regularly updated, etc. By doing this, the information extracted from it will be more useful and valuable. Improving data quality includes, among other things, ensuring that it is complete and reliable as well as maximising the frequency of updates and adding contextual information, metadata, etc.

One of the ways we can ensure a high degree of quality is to standardise similar data sets. This could be applied, for example, to municipal budgets or contracting information from different departments. Standardisation involves using the same fields of information and the same units of measurement, etc., so that the data can be compared and combined to extract knowledge, make improvements or solve problems in a specific field. In this regard, it is advisable to adopt international standards wherever possible, not only to facilitate the interoperability of data between the different Catalan administrations but also because it provides a larger pool of comparable data. This is particularly relevant for work in areas such as Big Data, which uses large volumes of data and is, therefore, facilitated by standardisation.

The more comprehensive the data, the more it can be exploited and the greater the knowledge and profitability that can be extracted from it. For this reason, we are prioritising the incorporation of historical references, georeferencing and the inclusion of categories that facilitate sectoral data analysis. For example, one of the obligations established in the Government's strategy is the incorporation of information relative to sex and gender in any data set that, by its very nature, is likely to contain that information. Data with a higher degree of disaggregation leads to more accurate analysis because it can be individualised or categorised (by gender or region, for example).

However, despite the immense efforts being made by our public administrations to publish large quantities of quality data, we must be aware that the act of opening data alone is worthless. This is not the objective. It is the subsequent processing of that data, its usage, analysis and understanding that generates value. Opening data is purely a means of achieving other goals, such as:

- Transparency and accountability. Clearly, from the moment data on subjects like public contracts, subsidy recipients and with whom the managers of public institutions hold meetings enters the public domain, citizens will have access to information that will allow them to demand we justify our actions. Citizens' loss of confidence in public institutions can only be addressed through transparency in the decisions and actions of those institutions, and the evaluation of their management.

- Encouraging citizen participation. Increasing the democratic quality of our country must, without a doubt, be achieved by increasing the participation of civil society in public decision-making. Society's collaboration in defining and implementing public policies cannot be limited to exercising its right to vote during the election of its representatives; it must be involved throughout the term of office. And it is these elected representatives who must provide the tools to enable that collaboration. For society to participate, it must have all the information related to the matter in hand at its disposal. Moreover, that information must be available in an intelligible format that allows for cross-examination and the extraction of additional data not provided by the Administration.

- Economic growth and boosting innovation. Open data has become a catalyst for

innovation. It contributes to the emergence of new businesses and creates jobs in the new information industry, from infomedia companies (intermediary organisations between the producers and consumers of information) to app developers. Furthermore, new professional roles such as data journalists (journalists specialising in research and analysis using open data), or “data scientists” are already appearing and seem to be in increasing demand in all organisations, including, of course, public administrations. The mission of this new profession is to extract knowledge from data, generate value and apply it to make improvements to an organisation. Concepts like big data, which generate fresh knowledge by combining large amounts of data from different sources, are also relevant.

- Supporting research. Open data leads to better quality research because it reduces the costs involved in data collection and facilitates data extraction through technology, which enables mass data processing. Furthermore, being able to publish the results of that research in an open data format facilitates knowledge sharing and prevents project repetition, saving resources as a consequence.

- Improving public services. With regard to our public administrations, open data can help us optimise the provision of public services, reduce costs and improve decision-making. Interoperability is just one example of the associated internal benefits; open data facilitates the sharing of public data between different entities without the need for data sharing agreements or the technological integration of different information systems. Open data will help to improve public services because it facilitates the detection of inefficiencies and the optimisation of internal processes. We must internalise the practices of problem-solving, planning and evaluating on the basis of objective and quantifiable data so that it becomes a part of our organisational DNA.

Because, as we said previously, the act of opening data alone has no direct benefit, it’s essential we empower citizens by providing them with knowledge around the usage of this instrument. The public sector must be ambitious and highly proactive in working towards this goal. The challenge is to convert data into information and information into knowledge, to extract maximum value from that data and to achieve maximum impact at all levels: social, economic, academic, etc. Our goal is to ensure the data we publish is useful and understandable. This is where the third pillar we mentioned at the beginning comes into play: encouraging reuse and promoting communicative and educational actions to disseminate knowledge and raise awareness among citizens, civil society, the business world and the administrations themselves of open data’s potential to create value.

We are living in the age of information, and there is enormous potential for all stakeholders – society, government, organisations, etc. – to become more transparent, responsible and ethical. Unfortunately, however, there is still a lack of knowledge around the full potential of using open data as an agent of change, and we are working hard to resolve this situation every day. Data, information and knowledge are power, and public institutions have a duty and a responsibility to enable that power for the benefit of all, not just the few.



Núria Espuny

Núria Espuny és directora general de Transparència i Dades Obertes del Departament d'Acció Exterior, Relacions Institucionals i Transparència de la Generalitat de Catalunya. És Enginyera Tècnica en Telecomunicacions i Enginyeria Electrònica per la Universitat Ramon Llull i té un postgrau en administració pública i societat de la informació per la Universitat Pompeu Fabra.