Open Data to Improve Sharing and Publication of Information between Public Administrations

Laura Gutiérrez Medina, José Luis Roda García, Guillermo García Juanes,
Alioth Rodríguez, Pedro González
ETSII, Universidad de La Laguna, Tenerife, Spain
jlroda@ull.edu.es

Summary: The Canary Islands receive 10 million tourists every year. Tourism represents a key sector for economic development in the Canaries. This work presents the benefits of open data usages in the tourism sector both in municipalities and in the island government. These public institutions have valuable information that must be shared: 600 hotels and apartments, 10,000 bars and restaurants, and more than 15,000 retail businesses. We are presenting an open data project to validate and publish such data across multiple administrations. The public sector can share their data sets to enrich the information and provide citizens and developers with a more accurate repository.

1. Introduction

The Canary Islands comprise an autonomous region in Spain. The population of the island of Tenerife is one million and each year it receives five million tourists. The public sector in the island is organized in one main administration called the *Cabildo de Tenerife* with different responsibilities at island level. The island is divided into thirty-one municipalities with very different population sizes, geography and economic characteristics.

In this work we present our experience in promoting open data in the context of the local public sector of the island. This project gives a new perspective to the public sector stakeholders that participate in open data strategy.

In 2012, an agreement between the Department of Education, Youth and Equality of the *Cabildo de Tenerife* and University of La Laguna for the promotion and development of Open Data strategy was established. The main goal was to open data sources related to mobility, training and employment in the context of University of La Laguna. During early stages of the project, it was extended to include other scopes and new stakeholders (tourism, transport and culture).

The initial step of the project was to convince the public administration participants of the convenience of providing citizens with the information they had collected or generated. The first meetings focused on spreading Open Data principles [1], European regulations and related national legislation. Specifically, in our case, the legal framework is the Spanish Law 37/2007, a state law on Re-use of Public Sector Information. Royal Decree 1495/2011 extends the previous law and establishes mandatory regulations for the general government administration for the publication of data from different areas in open format. The latter provides an element that we found very interesting for the implementation of this project: the *Re-use of Public Sector Information Plan,* the main objective of which was to promote, help and develop open data strategy.

This work presents the Canaries Open Data Project and the current cases of open data use by different administrations in the island. Thanks to open data strategy, these administrations can collaborate for validation and data sharing reliably and efficiently.

2. Canaries Open Data Project

The Canaries Open Data project is the first open data initiative in the Canary Islands. Its main objective is to promote, support and provide an open data web portal to citizens, developers and to public administration (http://www.opendatacanarias.es). The Canaries Open Data Project is led by the University of La Laguna (ULL) with funding from the *Cabildo de Tenerife*. The fact that the project is coordinated by the ULL permits greater independence and the ability to combine efforts between different administrations. Another key element in the ULL leading the project is the proximity to computer engineering students and researchers, which allows direct use of data through the development of web and mobile applications.

The following tasks we established:
- Identification of data providers, their data sets and formats.
- Requirements, development and loading data into the Canaries Open Data platform.
- Development of Web and Mobile prototypes and apps.
- Dissemination (i.e. competitions, media and conferences).

Once data providers were identified, we focused on the data quality that each organization was going to contribute in the first phase. On the one hand, we had the island administration of the *Cabildo de Tenerife*, with broad responsibilities in the island, which provided data related to tourism in the island: accommodation, restaurants and shops. Three municipal councils participated by providing data that complemented the information from the *Cabildo de Tenerife*. Also, public corporations such as *Turismo de Tenerife* or *Museos de Tenerife* participated by providing valuable information about their activities.

This task involved multiple meetings with each of the organizations. Each department described the structure of its organization and its responsibilities. Secondly, they presented us with the candidate data sets to publish on the Canaries Open Data platform. This was the point where coordination by the ULL played an important role. Each stakeholder worked with the ULL to provide useful data. The ULL had a global view of the data sets and detected duplicated, complementary and overlapping information between the providers. Each administration focused on their competencies but many details were of interest to two or more organizations. After reviewing the data in depth, errors, inconsistencies and duplications were detected. We informed the data providers, and in a second phase, they corrected their data sets. Finally, with the publication of the data on the Canaries Open Data platform, the different organizations could validate, compare and contrast the data sets between several administrations.

Each of the organizational units or institutions involved in the project was instructed in the need for the creation of the *Re-use of Public Sector Information Plan*. This Plan was proposed in the Royal Decree 1495/2011 which allows the general government administration one year to adapt to this law and to determine the four data sets that they are going to publish in the following six months. A Plan template has been published in [2]. We have adapted this

document to the goals of our project. The Plan contains the following items that each administration should describe and implement: a) Introduction of the Document, b) Background of the Administration, c) Definition of the Working Group, d) Identification and Selection of Reusable Information, e) Data Set Preparation Process, f) Conditions for Reuse, g) Designing and Management of Web Portal, h) Publication and Maintenance of Reusable Data, i) Measures for Assessment and Quality Improvement, j) Timeframe to Publish New Data Sets.

The ULL team again offered help and dedication to the development of the Re-use Plans for different administrations. Our support placed special emphasis on having well defined time limits, thinking about what data were going to be published, opening procedures and the state of their own data.

In parallel with these activities, the Canaries Open Data platform was developed based on customizing Drupal and CKAN platforms [3]. Several transformations of data set formats (csv, xml, rdf) and best practices documents for governments (e.g. data opening process, and how to develop APIs) were developed.

In June 2013, a public presentation of the platform was made to improve visibility. The idea was to encourage more administrations to take an interest in the project. In July 2013, a competition was held where several web applications and apps were developed using the platform data sets. The results were very exciting because on one side the project got great visibility, and on the other, the public sector employees of the participating administrations could see the huge potential of the open data initiative.

3. Case Studies: Tourism

The tourism industry accounts for over 70% of the economy of the Canary Islands. The *Cabildo de Tenerife* pays special attention to providing high quality tourist information. In the tourist information that we have opened on the platform, we highlight Tourist Information Offices, Accommodation, Bars and Restaurants, Shops, and Museum Events. There are other data sets on the platform not related to this work.

The following practical cases show the benefit of opening data for public administrations.

a) **Accommodation Data Set.**
   Both the public corporation *Turismo de Tenerife* [4] and thirty-one municipalities are collecting data for different types of accommodation. This represents about 500 hotels, apartments, rural hotels and houses, etc. The information that each municipality collects should be similar to the information held by *Turismo de Tenerife* which has island-wide accountabilities. The opening of data in this case allows the establishment and integration of information from various sources. The most advanced case is the Municipality of Arona which has also developed a CKAN based open data platform [5] with data in different formats including RDF. Therefore, *Turismo de Tenerife* and the Municipality of Arona can now access each other's data and enrich their data sources with information that was not previously available.

b) Restaurants Data Set.

This case is very similar to that above but in the field of bars, restaurants and cafés, etc. Around 10,000 of these businesses exist at this moment. We, again, have very similar data at both island administration department and municipality level. This is a clear case where integration of data sets would enrich each administration. Administrative information of the businesses is complemented with customer information, from different points of view.

c) Shops Data Set.

The Data Bank Department of the *Cabildo de Tenerife*, collects information from all retail stores that exist across the island. Municipalities also have to collect this data for their own administrative purposes (e.g. taxes, refuse collection, etc.). Although there is communication between the island departments and the municipalities, data updating is not very frequent (once a year, or less). With the publication of both data sets in our open data portal, all providers involved can efficiently share and update the information required. In this case, the municipalities of Arona and Tegueste have already offered their retail stores data, and the process of updating and improving it is now real time.

d) Museum Events Data Set.

Museos de Tenerife [6] is a decentralized administrative entity which was created by the *Cabildo de Tenerife* with the main mission of studying, cataloguing and preserving the island's cultural heritage. The different museums that make up the network comprise

important cultural spaces in the island that are visited not only by tourists but also by the citizens of Tenerife. In this case, we studied two data sets. On the one hand, the basic information held by the museums (address, location, timetables, web, email, phone) and on the other hand, the events held every day at the six museums (museum, event title, description, timetable, price). The first data set was a file in *CSV* format and the second data source was based on a *RSS* already publicly available offering details of all museum events.

Another benefit of releasing this data set is related to the case of *Turismo de Tenerife* mentioned above. By opening Museums' data sets, *Turismo de Tenerife* could automate their manual events updating process to integrate museum data into their database. Any other local council or island administration could, of course, also read events online from the Canaries Open Data platform. Having this information updated at

all times, *Turismo de Tenerife* or any other corporation could offer tourists more reliable information.

Three weeks after publishing these data in Canaries Open Data portal, two computer engineering students developed an app with all the information of the Museums of Tenerife. Its users not only include citizens and tourists, but museum staff that can also use the app to confirm the correct publishing of their events.

Currently, we are continuing to work on the project to offer new data sets related to transportation, environment, sports facilities, other events, etc. In every case, the government is the first to benefit from sharing their data sets and promoting new services, so that new companies can develop web portals and apps.

4. Benefits of the Canaries Open Data Strategy.

a) Opening public sector data is beneficial for citizens (transparency) and software developers (innovation) but also for administrations, as they can share data.

b) Using open data formats, public administrations can easily share and update data from various sources.

c) Greater inter-operability can be achieved when the same open data standards are used by different public administrations.

d) Customer service improvements. Data sets combined from several sources enrich the information shared and permit the development of new services for citizens. For example, the Canaries Open Data portal offers transportation data sets: bus and tram. In both cases, the companies responsible for these services (Metropolitan and TITSA) have developed the API to access routes, timetables and stops information. It is easy to relate this data to data about accommodation, restaurants and/or events, to offer alternatives for tourists to get around the island.

e) Open data strategy will create multiple business opportunities if the public sector continues to open their vast amounts of information.

References
[1] Open Data Principles. http://opengovdata.org/
[2] Re-use of Public Sector Information Plan. http://datos.gob.es/?q=node/1758 (in Spanish)
[3] Comprehensive Knowledge Archive Network. http://ckan.org
[4] Turismo de Tenerife. http://webtenerife.es
[5] Arona Open Data Portal. http://datos.arona.org
[6] Museos de Tenerife. http://www.museosdetenerife.org/