



# Open Data Strategies and Experiences to Improve Sharing and Publication of Public Sector Information

Laura María Gutiérrez Medina

*TARO Research Group, University of La Laguna, SPAIN, [lgutmed@gmail.com](mailto:lgutmed@gmail.com)*

José Luis Roda García

*TARO Research Group, University of La Laguna, SPAIN, [jlroda@ull.edu.es](mailto:jlroda@ull.edu.es)*

Guillermo García Juanes

*TARO Research Group, University of La Laguna, SPAIN, [ggjuanes@gmail.com](mailto:ggjuanes@gmail.com)*

Alioth Rodríguez Barrios

*TARO Research Group, University of La Laguna, SPAIN, [alioth.riguez@gmail.com](mailto:alioth.riguez@gmail.com)*

Pedro González Yanes

*TARO Research Group, University of La Laguna, SPAIN, [pgonyan@ull.edu.es](mailto:pgonyan@ull.edu.es)*

**Abstract:** *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 should be shared with other institutions: 600 hotels and apartments, 10,000 bars and restaurants, and more than 15,000 retail businesses. This article describes an open data project to validate and to publish such data across multiple administrations. The main benefits for the public sector are the improvement of the data quality and the interoperability between different administrations.*

**Keywords:** Open Data, Open Government Data, Public Sector Benefits, Tourism.

**Acknowledgement:** We would like to thank Cabildo de Tenerife, Turismo de Tenerife and Museos de Tenerife for their support and enthusiastic collaboration in their data process contribution. This work has been supported by the Share-PSI European project.

**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 the 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 to provide citizens with the information they had collected or generated. The first meetings focused on spreading the Open Government Data principles (<http://opengovdata.org/>), the 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. The 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 the open data strategy, these administrations can collaborate for validation and data sharing reliably and efficiently.

## 1. Canaries Open Data Project

The Canaries Open Data project (<http://www.opendatacanarias.es>) 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. 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 of 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 were established:

1. Identification of data providers, their data sets and formats.
2. Requirement, development and loading data into the Canaries Open Data platform.
3. Development of Web and Mobile prototypes and apps.
4. 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 the candidate data sets to be published 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 ("Plantilla para el desarrollo del Plan RISP", 2012). We have adapted this document to the goals of our project. The Plan contains the following items that each administration should describe and implement:

1. Introduction to the Document
2. Background to the Administration
3. Definition of the Working Group
4. Identification and Selection of Reusable Information
5. Data Set Preparation Process
6. Conditions for Reuse
7. Designing and Management of Web Portal
8. Publication and Maintenance of Reusable Data
9. Measures for Assessment and Quality Improvement
10. 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, what data sets 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 the Drupal (Corlosquet, S., Delbru, R., Clark, T., Polleres, A., & Decker, S, 2009) and CKAN platforms ([www.ckan.org](http://www.ckan.org)). In addition, several transformations of data set formats (CSV, XML and 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 (<http://www.opendatacanarias.es>). 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 (<http://www.opendatacanarias.es/apps>). 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.

## 2. Real Experiences in the Tourism Sector

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.

### 2.1. Accommodation Data Set

Both the public corporation *Turismo de Tenerife* (<http://www.webtenerife.com/>) and 31 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 (<http://datos.arona.org/>) 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.

## 2.2. Restaurants Data Set

This case is very similar to that above, but relates to the field of bars, restaurants and cafés, etc. Around 10,000 of these businesses exist at this moment. Again, we 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.

## 2.3. 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.

## 2.4. Museum Events Data Set

*Museos de Tenerife* (<http://www.museosdetenerife.org/>) 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 and phone) and on the other hand, the events held every day at the six museums (museum, event title, description, timetable and price). The first resource consists of a text file in CSV format. The second one is a syndicated RSS source already publicly available with the events of all the museums.

Another benefit of releasing this data set is related to the case of *Turismo de Tenerife* mentioned above. *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 and give tourists and citizens more reliable information.

Three weeks after publishing these data in the 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 also museum staff who can 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.



Figure 1: Screenshot of an app developed using the museums dataset.

### 3. Results and Benefits of the Open Data Strategy

During the development of the Canaries Open Data Portal several benefits have been achieved. Opening public sector data is beneficial for citizens (transparency) and software developers (innovation) but also for administrations, as they can share data (Ferrer-Sapena, A., Peset, F., & Aleixandre-Benavent, R., 2011).

By using open data formats, public administrations can easily share and update data from various sources. At present, different administrative services of the *Cabildo de Tenerife* often require data from other services. Due to the Canaries Open Data Project different administrative services are exploiting the potential of open data to share information with other administrations.

Greater inter-operability can be achieved when the same open data standards are used by different public administrations. We have collaborated with the *Cabildo of Tenerife* and municipalities to show that information interoperability can be achieved with open data. In this case we have worked with two specific examples: accommodation and restaurants.

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 (TITSA and Metropolitan) have developed an 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.

Open data strategy will create multiple business opportunities if the public sector continues to open their vast amounts of information (Alani, H., Dupplaw, D., Sheridan, J., O'Hara, K., Darlington, J., Shadbolt, N., & Tullo, C., 2007). Within the Canaries Open Data portal we have promoted through competitions the development of *apps* using museum events, accommodation and restaurants datasets.

## References

- Alani, H., Dupplaw, D., Sheridan, J., O'Hara, K., Darlington, J., Shadbolt, N., & Tullo, C. (2007). Unlocking the potential of public sector information with semantic web technology (pp. 708-721). Springer Berlin Heidelberg.
- Barta, R., Feilmayr, C., Pröll, B., Grün, C., & Werthner, H. (2009, June). Covering the semantic space of tourism: an approach based on modularized ontologies. In Proceedings of the 1st Workshop on Context, Information and Ontologies (p. 1). ACM.
- Bauer, F., & Kaltenböck, M. (2011). Linked Open Data: The Essentials. Edition mono/monochrom, Vienna.
- Bennett, D., & Harvey, A. (2009). Publishing open government data. W3C Working Draft. URL: <http://www.w3.org/TR/gov-data>.
- Bizer, C. (2009). The emerging web of linked data. *Intelligent Systems, IEEE*, 24(5), 87-92.
- César, C., & Lorenzo, S. (2010). Open government: gobierno abierto. Jaén, España: Algón Editores MMX, 2010.
- Corlosquet, S., Delbru, R., Clark, T., Polleres, A., & Decker, S. (2009). Produce and Consume Linked Data with Drupal! (pp. 763-778). Springer Berlin Heidelberg.
- EC European Commission. (2007). Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions. COM (2007), 780.
- Ferrer-Sapena, A., Peset, F., & Alexandre-Benavent, R. (2011). Acceso a los datos públicos y su reutilización: open data y open government. *El profesional de la información*, 20(3), 260-269.
- Fodor, O., & Werthner, H. (2005). Harmonise: a step toward an interoperable e-tourism marketplace. *International Journal of Electronic Commerce*, 9(2), 11-39.
- Gomez-Perez, A., Fernández-López, M., & Corcho-García, O. (2004). Ontological engineering. *Computing Reviews*, 45(8), 478-479.
- Hallo, M. A., Martínez-González, M. M., & de la Fuente Redondo, P. (2012). Las tecnologías de Linked Data y sus aplicaciones en el gobierno electrónico. *Scire: representación y organización del conocimiento*, 18(1), 49-61.
- Heath, T., & Bizer, C. (2011). Linked data: Evolving the web into a global data space. *Synthesis lectures on the semantic web: theory and technology*, 1(1), 1-136.
- Huijboom, N., & Van den Broek, T. (2011). Open data: an international comparison of strategies. *European journal of ePractice*, 12(1), 1-13.
- Janssen, K. (2011). The influence of the PSI directive on open government data: An overview of recent developments. *Government Information Quarterly*, 28(4), 446-456.
- Lassila, O., & Swick, R. R. (1999). Resource description framework (RDF) model and syntax specification.
- Ley 37/2007, de 16 de noviembre, sobre reutilización de la información del sector público. (2007). *Boletín Oficial Del Estado*, (276), 47160-47165. Retrieved October 2, 2014, from <http://www.boe.es/buscar/doc.php?id=BOE-A-2007-19814>
- Marcos-Martín, C., & Soriano-Maldonado, S. L. (2011). Reutilización de la información del sector público y open data en el contexto español y europeo. *Proyecto Aporta. El profesional de la información*, 20(3), 291-297.
- Plantilla para el desarrollo del "Plan RISP". (2012, June 29). Retrieved October 2, 2014, from <http://datos.gob.es/?q=node/1758>
- Prantner, K., Ding, Y., Luger, M., Yan, Z., & Herzog, C. (2007). Tourism ontology and semantic management system: state-of-the-arts analysis.
- Ramírez-Alujas, Á. (2011). Gobierno abierto y modernización de la gestión pública: tendencias actuales y el (inevitable) camino que viene. *Reflexiones semifinales. Revista Enfoques*, 9(15), 99-125.
- Ramírez-Alujas, Á. (2011). Open Government and Gobernanza Colaborativa: El (inevitable) camino hacia un nuevo paradigma de Gobierno y Gestión Pública. Estado del arte, desafíos y propuestas. In Ponencia presentada en el X Congreso de AECPA: "La política en la red", Murcia, del (Vol. 7).
- Real Decreto 1495/2011, de 24 de octubre, por el que se desarrolla la Ley 37/2007, de 16 de noviembre, sobre reutilización de la información del sector público, para el ámbito del sector público estatal. (2011). *Boletín Oficial Del Estado*, (269), 116296-116307. Retrieved October 2, 2014, from [http://www.boe.es/diario\\_boe/txt.php?id=BOE-A-2011-17560](http://www.boe.es/diario_boe/txt.php?id=BOE-A-2011-17560)
- The 8 Principles of Open Government Data (OpenGovData.org). (n.d.). Retrieved October 2, 2014, from <http://opengovdata.org/>
- Tu, S. (2009). Exploiting linked data to build web applications.
- Zeleti, F. A., Ojo, A., & Curry, E. (2014, June). Emerging business models for the open data industry: characterization and analysis. In Proceedings of the 15th Annual International Conference on Digital Government Research (pp. 215-226). ACM.

## About the Authors

### *Laura María Gutiérrez Medina*

Laura Gutiérrez is a contributor in the TARO Research Group at the University of La Laguna and a software engineering in an IT company. Her research interests include semantic technologies, Open Data, Linked Data and Big Data.

### *José Luis Roda García*

José Luis Roda García, PhD in Computer Science and Associate Professor in University of La Laguna (Tenerife, Spain). His main research topics include linked open data, semantic web, big data and model driven development. He is the project manager of the Canaries Open Data Portal. He works in open data projects to achieve interoperability and publication of public sector information. He is the local partner at University of La Laguna of the Share-PSI European network for the exchange of experience and ideas around implementing open data policies in the public sector.

### *Guillermo García Juanes*

Guillermo Garcia has worked in the Open Data project of University of La Laguna and always interested in topics related with Open Data, Linked Data and Software Engineering.

### *Alioth Rodríguez Barrios*

Alioth Rodríguez Barrios studied Computer Engineering at University of La Laguna and collaborates with Taro research group in the Linked Open Data topics.

### *Pedro González*

Pedro González works in the IT Faculty and collaborates with the TARO research group at the University of La Laguna. His research interests are models, business patterns, software engineering and open and linked data.