Open Data Assessment in Italian and Spanish Cities

Raffaele Sisto, Javier García López, José Manuel Paéz and Elena Maté Múgica

Abstract Citizens' attention to Open Data initiatives and Open Government policies is growing rapidly, but fragmentation in their implementation makes for a confusing approach. This research aims to bring structure to their assessment by analyzing 63 selected Spanish cities and 110 selected Italian cities and mapping their content. The analysis shows that (i) Open Data portals are not developed consistently; (ii) the various levels of public administration are not coordinated in their Open Data strategies; (iii) however, there are some good practices to be followed and underlined. Research into Open Data initiatives could benefit from previous examples regarding the success and failure factors of an Open Government. This paper highlights the main trends for an Open Data portal strategy in Spanish and Italian cities from the citizens' point of view. The basic content, structure of the websites, quality, accessibility, crosscutting nature and data visualization of the datasets have been assessed from the user's perspective. Currently, Spain has 115 open-data portals embedded in datos.gov.es (Government of Spain 2016), and Italy has 40 open-data portals at the city level. Open Data initiatives are the basis to achieve transparency, participation and collaboration, and also to establish the pillars for the integration of intelligent policies.

Keywords Smart city · Open government · Open data · Transparency
Urban level

R. Sisto (✉) · J. García López
Technical University of Madrid, Madrid, Spain
e-mail: raffaele.sisto@alumnos.upm.es

J. García López
e-mail: javier.garcialope@alumnos.upm.es

J. M. Paéz
Technologic of Monterrey, Monterrey, Mexico

E. M. Múgica
UCL, London, UK

A. Bisello et al. (eds.), *Smart and Sustainable Planning for Cities and Regions*,
Green Energy and Technology, https://doi.org/10.1007/978-3-319-75774-2_9

1 Introduction

In recent years, governments at the national, regional and local levels have acquired a new commitment towards horizontal strategies and performances, a trend known as Open Government. The number of Open Government initiatives has grown rapidly in the last ten years: Hundreds of government data catalogues have been published and over a million datasets have been released by governments around the world, spawning new businesses and social projects (Dawes and Parkhimovichc 2016). This idea of governance is based on policies that promote three main fields —Open Data, Transparency and Interaction with Citizens—by supporting the ongoing collaboration through bottom-up strategies. An Open Government is commonly seen as a driver of efficiency and a vehicle for increasing transparency, citizen participation and innovation in society (Jetzek et al. 2012).

This transition from passive to active citizenship needs information and communications technology (ICT) that is a powerful tool for triggering interactive, participatory and information-based urban environments (Garau et al. 2017). In fact, as shown in this study, public administrations, mainly municipalities, are increasingly opening specific websites dedicated to Open Government or any of its fields.

In a forecasting effort, Misuraca (2010) defined new government scenarios by 2030. They affirmed that ICT is changing the government system and defined four scenarios according to the level of transparency and openness and the level of integrated policy intelligence: Open Government, Leviathan Government, Privatized Government and Self-Service Governance. With regard to the Open Government, they defined:

An Open Government counts with a high level of both transparency and openness and integrated policy intelligence. Open Government will bring unprecedented access to knowledge, a shift of cognitive capacities, allows to personalized public services, increases the online engagement of citizens, new options or molecular democracy and improves risk management (Misuraca 2010).

Open Data, Transparency and Participation are not new concepts, they are as old as Democracy, but it is now, thanks to ICT advantages, when Government may afford a massive inclusion of citizens in the decision-making process (Calderon and Lorenzo 2010).

This Open Government scenario is considered the most likely model to be achieved. Thereby, following the Open Government principles, the vast majority of public administrations are trying to improve citizen information through openness, transparency and participation. In this context, the Open Data strategy has emerged for the promotion of free access to public-sector data, which can be reused by others to attract and serve citizens, businesses, and any other institution (García 2015), or to improve decision making and the quality of life (Rojas et al. 2017).

This paper aims to assess one of the most basic tools to continue on the path of the Open Government: Open Data. It is the basis of the two analyzed concepts—transparency and openness—and the support for integrated policy intelligence. As a

transparency tool for citizens, Open Data portals provide a framework to make government information available and to open the decision-making processes to citizens. Open-city data can help app developers, urban planners, and others understand the problems of a city and manage city services in ways that improve the quality of life and business prospects for its residents (Gurin 2014).

Governments generate and collect a vast quantity of data from many different domains (Janssen 2011). Urban space has become one of the leading data generators worldwide. Open Data is the true catalyst for Smart Services in the city (García 2015). Over recent years, various municipal departments have stored large amounts of data related to the city-government activity, but this information does not have a crosscutting nature. Open Data faces this issue and provides a return on public investment. "The release of this information is gaining increasing attention due to its potential to stimulate economic growth, support good governance and facilitate social innovation. Motivated by these benefits, several governments around the world have begun to include Open Data topics in their e-government strategies" (Rojas et al. 2017).

Data are a fundamental pillar of the Open Government strategy. It should be understood as an institutional, political and democratic innovation. Access to this information is not easy, and there is not a national consensus. In fact, many obstacles appear when these information sources are used: (i) data availability, (ii) willingness to use it, (iii) accessibility, usability and quality of data and (iv) the ability to collect, manage and use data (Stephenson et al. 2016).

In the case of Europe, the quality of Open Data varies depending on the country. In 2015, the European Data Bank conducted an assessment of where European countries stood with regard to Open Data. It showed that European public authorities had unequal levels of development in Open Data initiatives. Some countries are still in the process of launching a national Open Data portal, while other countries are already improving their portals and are starting new projects. As a conclusion, Italy and Spain are identified as the main trendsetters in the EU28+ Open Data Maturity clusters.

As required by the Digital Agenda for Europe, Italy has implemented policies for Open Data promotion through the Smart City National Observatory and the *Agenda Digitale Italiana*. Similarly, Spain approved in 2015, through AENOR (Spanish Association for Standardization and Certification), the standard UNE 178301 *Ciudades Inteligentes. Datos Abiertos* (*Open Data*). It is the first code that establishes a set of requirements for Open Data reuse in the public sector.

Therefore, the authors have chosen the cities from these countries in order to evaluate the main characteristics to be followed by the Open Data city managers.

This paper is structured as follows: Sect. 2 identifies the framework and describes of the fieldwork methodology. Section 3 shows the main results and a summary of all the analyzed fields; finally, Sects. 4 and 5 present the discussion and conclusions of this study.

2 Method

The main objectives of this study are: (i) to obtain a true and updated overview of the status of Open Data initiatives in the local administration of Spanish and Italian cities; (ii) to identify best practices within their efforts and strategies through an in-depth analysis of the characteristics of their official Open Data websites; (iii) to reveal gaps, such as the lack of standardization or homogeneous structure and the lack of a common strategy among various administrations to publish data openly; (iv) to assess the quality of Open Data regarding incorrect formats and useless or outdated data that do not meet the requirements in order to qualify as Open Data; (v) to help with the improvement of open-data strategies.

Thereby, it assesses the best portals in order to understand which data are more valuable and which are the most adequate techniques and technologies for both data scientists and citizen use.

As for the fieldwork methodology, the present study only analyzes official websites of local authorities. Central or regional government websites have not been considered. First of all, the Spanish and Italian provincial capitals and cities of over 200,000 inhabitants have been identified (the minimum size appropriately representative). After that, the Open Data portals of these cities have been assessed based on the following criteria: the presence of an official Open Data portal and/or a statistics website at the municipal level.

2.1 Open Data Assessment

After obtaining the state of the art of the preliminary Open Data portals, the quantity of datasets was analyzed. The structure and the general presentation of the portals of the selected cities were then analyzed as follows:

- Presentation: the access to datasets should be plain and functional, easy for any user to see. It should describe the mission and vision, the goals and the metrics to assess their achievements.
- Datasets: it should contain a search for keywords, filters, topics, tags, formats, and updated frequency to ensure maximum accessibility and navigability, plus a brief description of the formats used.
- Best Practices: the participation of third parties should be promoted and measured through the publication apps, best practices and examples of development based on datasets applications.
- Citizen Participation: it should include sections like Frequently Asked Questions, Contact and Tips, and provide a specific area for publishing applications developed by/for citizens.
- General Information: the terms, licenses and conditions of use of the datasets should be listed.

Table 1 Quantity of datasets

Range	Score	Cities (quantity of datasets)
0–49	0	Italy: Asti, Bergamo, Cagliari, Catania, Enna, Ferrara, Foggia, Forlì, Frosinone, Monza, Piacenza, Sassari, Siracusa, Verona; Spain: Albacete, Badalona, Cáceres, Ciudad Real, Cuenca, Girona, Lleida, Lugo, Santa Cruz de Tenerife, Tarragona
50–99	1	Italy: Ravenna (90), Livorno (88), Udine (80), Genova (74), Bari (57), Rimini (50), Napoli (50); Spain: Las Palmas (91), Sevilla (75), Santander (73)
100–199	2	Italy: Reggio di Calabria (160), Treviso (154), Venezia (152), Modena (150), Matera (138), Reggio Emilia (136), Pavia (119), Vicenza (118), Trento (106); Spain: Bilbao (174), Tarrasa (158), Sabadell (148), Cartagena (120), Valencia (113), Zaragoza (112), A Coruña (111)
200–299	3	Italy: Milano (291), Torino (263), Bolzano (212), Siena (207); Spain: Madrid (230), Pamplona (225)
More than 300	4	Italy: Bologna (1848), Firenze (1392), Roma (923), Pisa (606), Palermo (487), Lecce (344); Spain: Málaga (633), Gijón (564), Barcelona (331)

Table 1 shows the results of the first steps of the analysis: the quantity of datasets and the structure of the websites (Fig. 1).

The cities excluded in these two steps (20 Italian cities and 14 Spanish cities) are those with less than 50 datasets and those without an adequate website structure considered as a suitable Open Data portal (less than six points).

Finally, the 11 finalist Spanish cities and the 20 finalist Italian cities were fully assessed with regard to the following four dimensions: content quality, content accessibility, crosscutting nature of the content and content visualization.

2.1.1 Content Quality

In order to evaluate how the content of the dataset can be used and how reliable the data is, the quality of the data has been assessed with regard to the following criteria:

- The data available is "open source": the license of datasets must comply with the concept of Open Definition (reuse, redistribution and free use are permitted).
- The data available is free: a payment of a fee is not necessary in order to access to the datasets.
- The data available comes from an official source: the datasets must come from a government source or an officially recognized source. Otherwise, datasets are not considered reliable.
- A format based on "the 5-star Open Data" (Berners-Lee and O'Reilly 2009): it is a one-to-five-star scale that attempts to define the degree of reuse and openness of data depending on its format. Level 3 is the lowest for a dataset to be considered "open".

Cities selected	Presentation	Finder	Filters	Formats description	Best Practices	Data requests	Ideas and results	Terms of uses	Contact	Legal warning	Total
Lecce	★	★	★	★	★	★	★	★	★	★	10
Milano	★	★	★	★	★	★	★	★	★	★	10
Reggio di Calabria	★	★	★	★	★	★	★	★	★	★	10
Reggio Emilia	★	★	★	★	★	★	★	★	★	★	10
Treviso	★	★	★	★	★	★	★	★	★	★	10
Vicenza	★	★	★	★	★	★	★	★	★	★	10
Pisa	★	★	★	★	★	★	☆	★	★	★	9
Siena	★	★	★	★	★	★	☆	★	★	★	9
Torino	★	★	★	★	★	☆	★	★	★	★	8
Bolzano	★	★	★	★	☆	★	☆	★	★	★	8
Matera	☆	★	★	★	★	★	☆	★	★	★	8
Bologna	☆	★	★	★	☆	★	★	☆	★	★	7
Venezia	★	☆	★	★	☆	★	☆	★	★	★	7
Ravenna	★	★	★	★	☆	☆	☆	★	★	★	7
Rimini	★	★	★	★	☆	☆	☆	★	★	★	7
Palermo	★	★	☆	★	★	☆	☆	★	★	★	7
Firenze	☆	★	☆	☆	★	★	☆	★	★	★	6
Roma	☆	★	☆	★	☆	★	★	★	★	★	6
Pavia	★	★	★	☆	☆	☆	☆	★	★	★	6
Udine	★	☆	☆	★	☆	★	☆	★	★	★	6
Modena	★	★	★	☆	☆	☆	☆	★	☆	★	5
Trento	★	★	☆	☆	☆	☆	☆	★	★	★	5
Livorno	☆	★	☆	☆	☆	★	☆	★	★	★	4
Bari	★	★	★	☆	☆	☆	☆	★	★	☆	4
Genova	☆	★	★	☆	☆	☆	☆	★	★	☆	3
Napoli	★	☆	☆	★	☆	☆	★	☆	☆	☆	3

Cities selected	Presentation	Finder	Filters	Formats description	Best Practices	Data requests	Ideas and results	Terms of uses	Contact	Legal warning	Total
Zaragoza	★	★	★	★	★	★	★	★	★	★	10
Madrid	★	★	★	★	★	★	★	★	★	★	10
Bilbao	★	★	★	★	★	★	★	★	★	★	10
Sevilla	★	★	★	☆	★	★	★	★	★	★	9
Santander	★	★	★	☆	★	★	★	★	★	★	9
Barcelona	★	★	★	☆	★	★	★	★	★	★	9
Valencia	★	★	★	★	★	★	☆	★	★	★	9
Las Palmas	★	★	★	☆	★	★	★	★	★	★	9
Malaga	★	★	★	☆	★	★	☆	★	★	★	8
Gijón	★	★	★	★	★	☆	☆	★	★	★	8
Tarrasa	★	★	★	★	☆	☆	☆	★	★	★	7
Sabadell	☆	★	☆	★	☆	☆	☆	★	★	★	5
A Coruña	☆	★	★	☆	☆	☆	☆	★	★	★	5
Cartagena	☆	★	☆	☆	★	★	☆	★	★	★	5
Pamplona	☆	★	★	☆	☆	☆	☆	☆	★	★	4

Fig. 1 Website structure. Each star equals one point. Cities with less than six stars were excluded

Level 1—Open License: format files such as PDF, DOC, etc.
Level 2—Readable: format files such as XLS, etc.
Level 3—Open Format: format files such as CSV, XLM, etc.
Level 4—Resource Format: format files such as HTTP, RDF etc.
Level 5—Linked Data: published data linked and pointed to other data sources and contexts.

- The data available is updated: the updating of the datasets is evaluated taking into account metadata, file name description, data tags and date of publication.

2.1.2 Content Accessibility

The user accessibility to the data included in the datasets were assessed with regard to the following five criteria:

- Download grouping: all datasets can be downloaded, at once and in single file (RDF, XLS, etc.).
- SPARQL/API Service available to be used by robots.

- Languages: the availability of different languages has been assessed. It has been graded depending on the number of languages. The native language has not been considered.
- City-council website: a direct link to the portal from the city-council website has been considered.
- Online and downloadable data: if the data can only be sent by mail, it is not considered available.

2.1.3 Cross-Cutting Nature of the Content

The distributed datasets have been evaluated according to the government areas and their association, using the classification from "Ranking of European medium-sized cities" (Giffinger and Fertner 2007) as follows:

- Economy: commerce, finances, employment, taxes, industry, tourism and housing.
- Mobility: transport and urban infrastructure.
- Environment: energy and environment.
- Citizens: demography and education.
- Quality of life: sciences, technology, culture, leisure, safety, health, sport, society and wellness.
- Governance: law, justice, public sector.

2.1.4 Content Visualization

An interactive platform or an online dashboard was used to analyze the ability to visualize each dataset before downloading it, taking into consideration the following four aspects:

- Online Dashboard (four points): if there is an online dashboard available for several datasets.
- Data Pre-visualization (three points): if several datasets can be previewed before downloading.
- Readable Format (two points): if the dataset formats are accessible for non-professional users (file formats such as xls, XML etc.). Only Open-source formats have been considered.
- Download Average (one point): if the average download of datasets is displayed (Fig. 2).

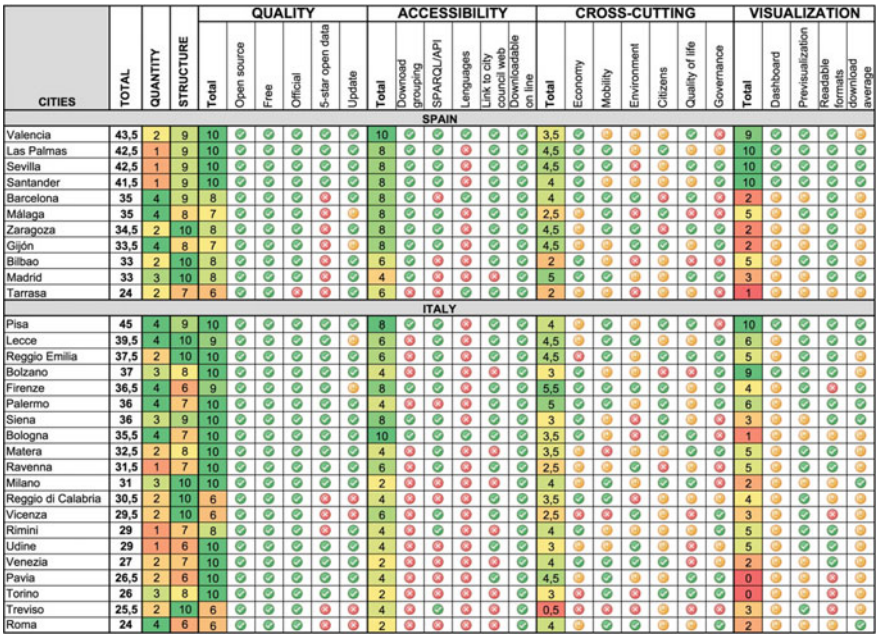

Fig. 2 Total results for Spain and Italy. Green circle = maximum score; yellow circle = medium score; red circles = minimum score (for more information about scoring, see Sect. 2)

3 Results

Table 2 shows the results for each area analyzed. Valencia and Pisa are the best-valued cities and have an excellent score in all areas. Virtually all the portals respect the basic concepts (quality and content of website). On the contrary, there is a huge difference between the Spanish and the Italian portals in terms of accessibility: more than half of the Italian cities do not guarantee the minimum requirements to be considered accessible. Málaga, Tarrasa and Treviso do not include a variety of datasets; they focus the data mainly in one or two domains instead. With regard to the visualization, there is no uniformity in the display method.

Table 2 The average percentage of the datasets for every area and a range have been established

Economy	Mobility	Environment	Citizens	QoL	Governance	Score
0–4.7%	0–6.0%	0–2.4%	0–8.5%	0–5.8	0–5.3	0
4.7–9.4	6.0–11.9	2.7–5.3	8.5–17.8	5.8–11.5	5.3–10.6	0.5
9.4–21.2	11.9–26.9	5.3–12.0	17.8–40.1	11.5–26.0	10.6–23.9	1
21.2–42.3	26.9–53.7	12.0–24.0	40.1–80.1	26.0–51.9	23.9–47.7	0.5
+42.3	+53.7	+24.0	+80.1	+51.9	+47.7	0

4 Discussion

The main impression is that Spanish cities have decided to open their data in a very heterogeneous way. There are many gaps in both form and content because they have not followed a common strategy.

On the one hand, there are cities with transparent websites that have not opened their data (e.g., Murcia or Ciudad Real). Other cities that have a specific website but have content that is limited, irrelevant, unclear or does not respond to what is considered open data (e.g., Tarrasa and Ciudad Real only have jpg or pdf files respectively). On the other hand, cities like Pisa, Madrid and Valencia, among others, are as advanced as the North American ones (e.g., San Francisco, CA).

A vast majority of the analyzed municipal websites have direct links to any specific fields related to Open Government, but only 36.5% of Spanish cities (23/63) and 36.3% of Italian cities (40/110) have an official Open Data website.

Several autonomous governments, such as Aragón or Castilla y León in Spain and Veneto or Trentino-Alto Adige in Italy, have grouped their regional Open Data portals instead of dividing them into categories (city or province).

With regard to format types, 33 different types have been found. Ten of them represent 83% of all published datasets. The three most used formats by the 11 selected cities are CSV (24.2%), XML (11.7%) and JSON (11.4%). The formats used are mainly focused on promoting their interoperability.

5 Conclusion

This study reveals a lack of an in-depth analysis of the Open Data benefits for an Open Government strategy. The two major uses of Open Data portals are not properly represented in the majority of cities evaluated, namely:

- Citizens' empowerment and transparency. The main issue is the lack of a visualization and organization standard. None of the examples includes an easy module where citizens can ask for relevant information.
- Data business models. No analysis of the data value is evaluated in any example before deciding which data will be published.

Each administration level (national, regional or local) has begun to develop Open Government strategies with various scopes and methods. This is an ineffective approach due to their distribution of responsibilities and to some objections in the dissemination of certain information that could be understood as compromising. All this has led to a heterogeneous state of the art in content and form. Below are summarized some opportunities for improvement and also recommendations to achieve the principles described in the previous code of good governance:

- A common strategy among administrations: In addition to the regulatory framework, it would be advisable to standardize Open Government strategies and automatically share data among them.
- Municipal Strategic Plan for Open Government: The real needs of citizens should be prioritized during the boot process for transparency, openness and participation initiatives.
- Data Qualification: Due to the lack of commitment to data quality, data should be preselected o that it can be opened without becoming a useless mess. This is needed in order to standardize the format, structure and scope of Open Data initiatives and to share common principles: open, standardized, interoperable, linkable, machine-readable and non-proprietary formats.
- Encourage Citizen Participation: while progress towards municipal-management transparency has been shown, citizen participation must be promoted through advertising campaigns, grants, competitions, fairs and/or associations that use municipal resources in order to encourage innovation and entrepreneurship.
- Open Data Indicators: it is advisable to keep track through performance metrics that monitor the commitment of the stakeholders and to identify potential gaps and opportunities of improvement in Open Government policies.
- Open Data Dashboard: It is recommended to show the most relevant data through a dashboard. It should be free, accessible, online, in real time (when possible) and linked to the transparency portal.

Finally, since Spain and Italy are the European leading countries within the Open Data Maturity initiatives (European Data Portal 2015), the methodology used in this analysis aims to be applicable and valid in other contexts and may be useful to analyze Open Data initiatives in European cities.

Acknowledgements The authors take this opportunity to gratefully acknowledge the assistance and contributions of Dr. Sergio Ramos, Dr. José Maria Lapuerta, Javier Dorao, Estefanía Rueda and Daniel Martinez. And last, but not least, we would like to thank the SSPCR2017 attendees for their useful suggestions during the conference.

References

AgiD. (2015). *Agenda Digitale Italiana.* Accessed at: http://www.agid.gov.it [16-07-2017].

Berners-Lee, T., & O'Reilly, T. (2009). A conversation with Tim Berners-Lee. In *Web 2.0 Summit 09*.

Calderon, C., & Lorenzo, S. (2010). *Open government. Gobierno Abierto*. Algón Editores.

Dawes, S., Vidiasova L., & Parkhimovichc, O. (2016, January). Planning and designing open government data programs: An ecosystem approach. *Government Information Quarterly, 33*(1), 15–27.

EU Open Data Portal. Accessed at: https://open-data.europa.eu/ [15-12-2016].

European Data Portal. (2015). *Open data maturity in Europe. Insights into the European state of play*. European Union.

European Data Portal. Accessed at: http://www.europeandataportal.eu/ [10-12-2016].

Garau, C., Balletto, G., & Mundula, L. (2017). A critical reflection on smart governance in Italy: Definition and challenges for a sustainable urban regeneration. In *Smart and sustainable planning for cities and regions* (pp. 235–250). Berlin: Springer International Publishing.

García, M. (2015). *New businesses around open data, smart cities and fiware*. European Public Sector Information Platform. Topic Report 2015/04.

Giffinger, R., & Fertner, C. (2007). *City-ranking of European medium-sized cities* (pp. 1–12). Vienna UT: Centre of Regional Science.

Global right to information rating. Centre for Law and Democracy. Accessed at: http://www.rti-rating.org/ [16-06-2016].

Gurin, J. (2014). *How open data is transforming city life*. New York: FORBES.

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.

Jetzek, T., Avital, M., & Bjorn-Andersen, N. (2012). The value of open government data: A strategic analysis framework. In *SIG eGovernment Pre-ICIS Workshop 2012*.

Misuraca, G., et al. (2010). *Environmental digital Europe 2030: Scenarios design on future ICT for governance and policy modelling*. Seville: Joint Research Centre Institute for prospective technological advancements.

Municipal Transparency Index ITA. Accessed at: http://transparencia.org.es/ [02-07-2016].

OGDP. (2007). *Open government data principles: 8 principles of open government data*. Accessed at: http://www.opengovdata.org/home/8principles [15-04-2016].

Open data initiatives map. CTIC Foundation. Accessed at: http://www.fundacionctic.org/. [18-04-2016].

Rojas, L. A. R., Lovelle, J. M. C., Bermúdez, G. M. T., Montenegro, C. E., de Ory, E. G., & Crespo, R. A. G. (2017). Metamodel to support decision-making from open government data. *Journal of Ambient Intelligence and Humanized Computing,* 1–11.

Spanish Government Transparency Portal. Accessed at: http://transparencia.gob.es/ [22-01-2017].

Stephenson, P. J., Bowles-Newark, N., Regan, E., Stanwell-Smith, D., Diagana, M., Höft, R., et al. (2016). *Unblocking the flow of biodiversity data for decision-making in Africa. Biological Conservation*, *213*(2017), 335–340.

The open data handbook. Accessed at: http://opendatahandbook.org/ [25-04-2016].

U.S. General Services Administration. Accessed at: https://www.data.gov/ [03-03-2016].

Each official Open Government website of the assessed Spanish and Italian cities was visited during February and May 2016.