



Open data in the Italian Government: the experience of the City of Florence

Giovanni Menduni, Gianluca Vannuccini,
Giacomo Innocenti

The City of Florence, Italy, has been carrying out a large data integration and consolidation process since 2003, that was even more improved and targeted to an open and federated approach during 2009. Results of such a process are now being exploited in several current hot eGovernment fields, among which business intelligence, and, more recently, the open data movement. During the year 2011, the City of Florence lead an internal structured assessment process in which each department named an open data referee, and was called to analyze which available public data stores were eligible to be opened up in a suitable website section. It is worth noticing the strategic choice made by City of Florence, that decided to publish, when possible, information at its atomic level, in order to provide raw data, which can be directly analyzed by the final users. This approach has made the open data particularly suited for business intelligence purposes, but on the other hand the hidden informative richness makes such documents harder to read without proper inquiring tools. To this end, a child project has born. Its aim is to provide the common citizen with simple visual and graphical tools, designed to investigate the data via comparisons defined by



the user through parametric selectors. At the same time, the best practices to publish on the internet the open data have been studied, taking into account the different needs which can characterized the possible users. Based on the results of this research, it has been decided to export the open data made of collections of single electronic documents according to open and free formats, in order to not compromise the user experience because of the need of proprietary viewers and reading tools. Among all the possible choices, particular attention has been put on the more popular formats; specifically, for data directly deriving from database tables the CSV formats has been preferred to the others, while for geospatial data the choice has fallen to the KMZ format. City of Florence has also decided to encounter the needs of the local industry by starting an other project for the publication of its data as services. According to this paradigm the information is provided to the user according to parameters that the latter can specify at the moment of the request. This approach is particularly fruitful for software applications, because it reduces both the computational burden and the weight of the data transmission. To this aim the City of Florence has started a series of dialogs with independent developers, which have led to the production of several software application for mobile phones based of the open data. At first, a number of "apps" for Android and iOS smart phones has been realized by embedding the information directly into the software; then, new apps accessing the open data via specific online services has been designed and they are actually in testing phase. In the above framework, the City of Florence has also decided to used the data-as-service paradigm to federate its collections with the ones provided by other subjects. In this respect the Linked Data approaches been pursued, with particular attention to the W3C star-rating scheme suggested by Tim Berners-Lee, From a technical point of view, the linked data are made of rich entities

interconnected via semantic properties and the common method for their representation is a list of clauses describing both the features of each entity and their linking properties. To this aim it is used the RDF language and its derivations along with a specific storage system. Even though specific software and tools exist to handle these elements, a preliminary analysis drove to the adoption and experimentation of an alternative solution consisting of a wrapper that maps the data in a DBMS into an RDF graph at run-time. This choice has been preferred in order to not duplicate the open data with another storage system. Hence, the D2R server open source project has been choice as a basis for DBMS to RDF mapping. The first three datasets that were mapped in RDF are the municipal road network naming system, the corresponding house number system, and the rank of streets with more street-accidents. The first open data section of the city of Florence website was published in October 2011, to which followed the publication of RDF scheme, and a data-as-service connector, i.e. a SPARQL endpoint, for the three above datasets. In February 2012, a brand-new specific open data portal was published, whose main sections were: the dataset catalog; the "open data for all" section, with human-readable data visualizations helping the user to experience simple Business Intelligence tools; the linked data section; and, eventually, the GeoPortal section, giving access to geospatial data through standard catalog, storage and web portal, which are based on the open source GeoNetwork and GeoServer projects. Inside the open data project, the City of Florence has promoted the citizen participation by means of communication tools as the social networks. In particular, a specific hashtag #opendatafirenze was used for citizens engagement via Twitter, and for the City of Florence as a channel to publish the daily news. Since the new website launch, in February 2012, more than 120 tweets with this hashtag have been registered, and the "daily

dataset” is twitted every working day.¹ The main efforts of the City of Florence on this field are now particularly focused on the improvement of the RDF-mapped portion of the whole data store (the museums dataset was recently published), and on the improvement of the adopted dictionaries. Indeed, due to the lack of specific and easy-to-use semantic standards for public administration, we adopted a home-made dictionary to model the above RDF-mapped datasets. The early draft of the dictionary was intrinsically linked to the nature of the open data; however, this is not the best practice to define a semantic ontology. Hence, a further development has led to define a new dictionary, this time considering all the opposing needs: on the one hand, to be completely fruited, its logic has to be close to the data itself; on the other hand, it has to be sufficiently generic in order to be re-used in several different contexts. To this aim the new release of the dictionary propose has been internally organized in a deep structure, where the upper layers provide the generic property, while the lower ones are devoted to describe the details of the considered data. In this way, a user can adapt the dictionary by over-writing only the lower layers, when needed. To simplify this approach, the dictionary itself has been divided into a series of connected sub-dictionaries. Even though the related ontology is not providing yet a complete coverage of the Public Government needs, hopefully further collaborations with other public bodies (such as central national bodies, for Government ontology standardization) will better off the dictionaries.

¹<https://twitter.com/i/#!/search/?q=%23opendataFirenze&src=hash>.

GIOVANNI MENDUNI, Area di Coordinamento Programmazione, Sostenibilità e Innovazione - Comune di Firenze.

giovanni.menduni@polimi.it

GIANLUCA VANNUCCINI, Servizio Sviluppo Infrastrutture Tecnologiche - Comune di Firenze.

g.vannuccini@comune.fi.it

GIACOMO INNOCENTI, Dipartimento di sistemi e informatica - Università degli Studi di Firenze.

giacomo.innocenti@unifi.it

Menduni, G., G. Vannuccini, G. Innocenti. "Open data in the Italian Government: the experience of the City of Florence". *JLIS.it*. Vol. 4, n. 1 (Gennaio/January 2012): Art. #5500, p. 435–440. DOI: [10.4403/jlis.it-5500](https://doi.org/10.4403/jlis.it-5500). Web.

ABSTRACT: The City of Florence, Italy, has been carrying out a large data integration and consolidation process since 2003, that was even more improved and targeted to an open and federated approach during 2009. Results of such a process are now being exploited in several current hot eGovernment fields, among which business intelligence, and, more recently, the open data movement. During the year 2011, the City of Florence lead an internal structured assessment process in which each department named an open data referee, and was called to analyse which available public datastores were eligible to be opened up in a suitable website section. The first Open Data section of the City of Florence website was published in October 2011, to which followed the publication of RDF scheme, and SPARQL endpoint for the three above datasets. In February 2012, a brand-new specific open data portal was published, whose main sections were the dataset catalogue, the "Open data for all" section, with human-readable data visualisations, the Linked Data section, and the GeoPortal section, giving access to geospatial data through a standard catalogue and a portal which are based on the open source GeoNetwork and GeoServer projects.

KEYWORDS: City of Florence; Open data; eGovernment

Submitted: 2012-04-25

Accepted:

Published: 2012-05-16

