



SmartOpenData



This project has received funding from the European Union's Seventh Programme for research, technological development and demonstration under grant agreement No 603824.

PROJECT PERIODIC REPORT

Publishable Summary

Grant Agreement number: 603824

Project acronym: SmartOpenData

Project title: Linked Open Data for environment protection in Smart Regions

Funding Scheme: FP7 - THEME [ENV.2013.6.5-3 ENV.2013.6.5-3]

Date of latest version of Annex I against which the assessment will be made: 4th September, 2015

Periodic report: 1st 2nd 3rd 4th

Period covered: from M13 (November 2014) to M24 (October 2015)

Mariano Navarro de la Cruz, R&D & Innovation unit ICT manager, TRAGSA Group

Tel: + 34 91 322 65 21

Fax: + 34 91 322 53 00

E-mail: mnc@tragsa.es

Project website¹ address: www.smartopendata.eu

¹ The home page of the website should contain the generic European flag and the FP7 logo which are available in electronic format at the Europa website (logo of the European flag: http://europa.eu/abc/symbols/emblem/index_en.htm logo of the 7th FP: http://ec.europa.eu/research/fp7/index_en.cfm?pg=logos). The area of activity of the project should also be mentioned.

Declaration by the scientific representative of the project coordinator

I, as scientific representative of the coordinator of this project and in line with the obligations as stated in Article II.2.3 of the Grant Agreement declare that:

- The attached periodic report represents an accurate description of the work carried out in this project for this reporting period;
- The project (tick as appropriate)²:
 - has fully achieved its objectives and technical goals for the period;
 - has achieved most of its objectives and technical goals for the period with relatively minor deviations.
 - has failed to achieve critical objectives and/or is not at all on schedule.
- The public website, if applicable
 - is up to date
 - is not up to date
- To my best knowledge, the financial statements which are being submitted as part of this report are in line with the actual work carried out and are consistent with the report on the resources used for the project (section 3.4) and if applicable with the certificate on financial statement.
- All beneficiaries, in particular non-profit public bodies, secondary and higher education establishments, research organisations and SMEs, have declared to have verified their legal status. Any changes have been reported under section 3.2.3 (Project Management) in accordance with Article II.3.f of the Grant Agreement.

Name of scientific representative of the Coordinator: Mariano Navarro de la Cruz

Date: 26/ 01 / 2016

For most of the projects, the signature of this declaration could be done directly via the IT reporting tool through an adapted IT mechanism and in that case, no signed paper form needs to be sent

² If either of these boxes below is ticked, the report should reflect these and any remedial actions taken.

Table of Contents

Publishable Summary.....	4
Summary description	4
Work performed and main results	6
Expected final results and use	9

Publishable Summary

Summary description

The SmartOpenData project has completed, during this second period, the creation of a Linked Open Data -LOD- infrastructure extended throughout Europe. This research infrastructure has five different pilots as one of its most visible outcome. Nevertheless, those tools and applications are well grounded on several data modelling and data harmonization processes.

The SmartOpenData information environment has been fed, in most cases, by public and freely available existing sources for biodiversity and environment protection and research in rural and European protect areas, National Parks and tourist locations. However, SmartOpendata, has gone further in some specific cases, by publishing datasets that are not completely public or freely available for Citizenship.

In this context, SmartOpenData has reused and recycled existing information from several sources such as public open data portals like GEOSS/GEO Portal, INSPIRE and voluntary data sources like OpenStreetMap.

Accordingly, the SmartOpenData infrastructure is based, in first instance, on existing software applications and datasets. Nevertheless, in this second iteration, the project has also developed its own tools and has made public several new datasets using completely innovative publication tools and techniques, at least in regard to geospatial information systems.

Given the asymmetric situation of public data registries and publication across Europe, SmartOpenData, within its scope, has carried out some general tasks in regards to information processing and publication. In each project scenarios, the required datasets and the owners of that information have been sought out. As the original formats were found to be various and inhomogeneous, SmartOpenData partners have selected the most suitable tools for refining, transforming and publishing the data in each specific case or, in some cases, they have developed completely new applications.

However, in context of LOD, data publication is just one of the results that could be obtained. Final users demand more than just the information and data and, consequently, the project has made publically available all of the data models and ontologies that accurately describe the generated datasets that are published as LOD.

Therefore, SmartOpenData has defined mechanisms and strategies for searching, acquiring, adapting and publishing Open Data provided by existing sources regarding biodiversity and environment protection in rural and European protected areas and National Parks. In this second period of the project, the LOD obtained thereby has been used to solve specific semantic queries.

As a secondary, but no less important outcome, in addition to the technical aspects and results SmartOpenData has helped to reduce the gap between the geospatial community and Semantic Web movement led by international standards bodies and universities. Indeed, SmartOpenData has focused on how LOD can be explained to and disseminated into the geospatial community and applied generally to spatial data resources.

At this time, there are several issues that have already been properly answered regarding GIS systems and datasets, but there are still open questions in the context of spatial data as LOD. For

example, small geometries encoding or the implementation of topological functions have implemented correctly but those processes could be improved.

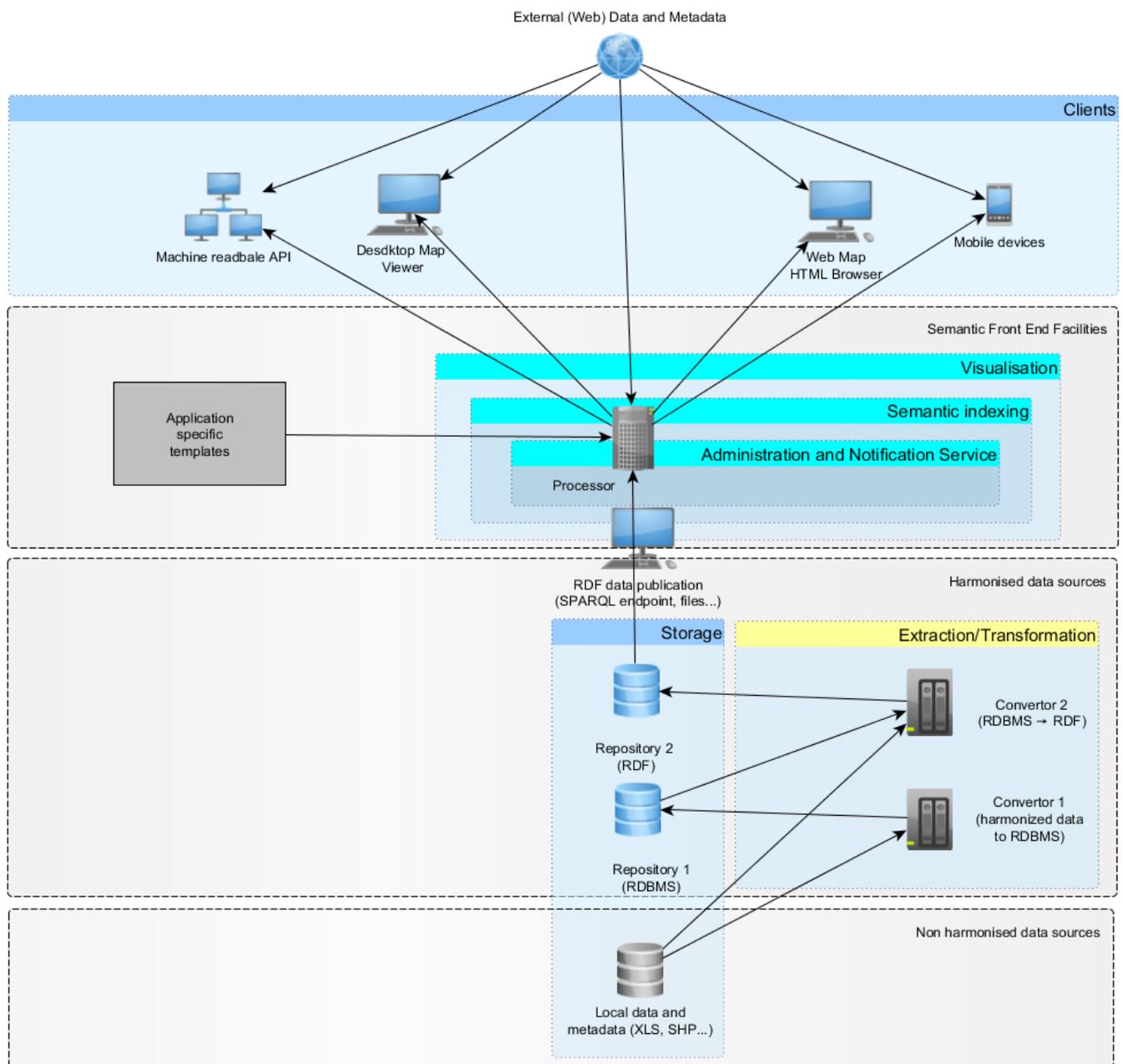
The vision of the SmartOpenData project in this second period has been widely confirmed: there are many different environmental information sources and their level of openness is varied. And, as a direct consequence, the economic value of the datasets can be greatly improved through its wide public exposition in a proper way.

As a final objective, SmartOpenData contacted several SMEs and stakeholders to offer them the results of the project, with the conviction that the power of Linked Open Data will foster innovation within the environmental sector.

Work performed and main results

The first 6 months of the project were focused on the analysis and definition of clear use cases that explain properly the power of Linked Open Data tools inside the context of geospatial environmental and biodiversity data. The required datasets were requested or obtained from several sources.

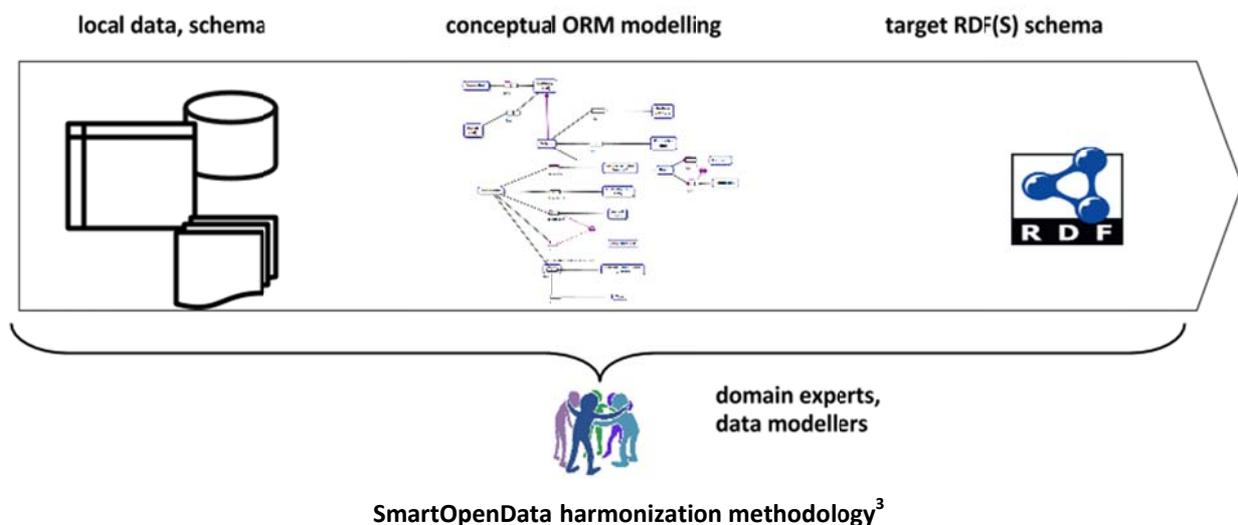
During the first year, the project defined a complete list of requirements that served as the starting point for pilots' definitions. Also, a complete study regarding semantic tools and architectures that could be used in geospatial environment was carried out. Based on these, the common SmartOpenData architecture was defined.



SmartOpenData general architecture

As one of the main foundations of the project, during this second period, SmartOpenData completed the definition of a powerful and flexible vocabulary (already published by W3C at <http://www.w3.org/2015/03/inspire/smod#>) that covers the central core of common issues of the pilots. For this SmartOpenData data model, and as far as possible, the project has used INSPIRE as the basis for the data structures in each pilot. Hereafter, using this central core, the pilots extended the SmartOpenData vocabulary to take into account their own singularities.

Once the data model was defined, a general data harmonization process was carried out jointly by partners responsible for pilots and technical partners. The final visible result of this harmonization process is publication of the Linked Open Datasets, and no less important, a conceptual modelling cycle supported by domain experts and data modellers discussions lies behind this.



On this LOD publication layer, software tools developed by the project deliver the full potential of linked information. Those tools have been:

SIREn indexing infrastructure as a highly scalable open-source full-text search engine especially suited for nested and schema-less data.

Sefarad is a web application developed to explore linked data by executing SPARQL queries at a chosen endpoint without writing code. Thus, it provides a semantic front-end to Linked Open Data that allows the user to configure his/her own dashboard with many different widgets to visualize, explore and analyse graphically the different characteristics, attributes and relationships of the queried data.

The **Administration and Notification Service** provides common facilities for exploiting environmental data. This refers to the fact that many solutions to the exploitation of environmental data are not interoperable, and any changes to the data source stops these solutions working. Also, this service improves environmental data searchability thanks to integration of big data infrastructure for structured and semi-structured search facilities.

The SmartOpenData pilots, as showcases of the tasks developed through the previously mentioned processes, demonstrate data modelling, data processing, data publication and technical tools. The first and final iterations of pilots, accomplished during the second period, has

³ ORM stands for Object-Role Modelling, one of the methodologies used in Data Harmonization process

shown practical examples of the SmartOpenData architecture and implementation of the SmartOpenData data models.

In this second period, the pilots, applications, tools and infrastructure were evaluated from the user groups' perspective. The goals of this evaluation have been to assess the quality of the technical results, verify whether they meet users' specified requirements and to collect feedback on the developed tools via interaction with the internal and external user groups represented by the relevant stakeholders. In this context, two evaluation phases were carried out in the project: an internal evaluation, involving internal user groups representing participants from the main project partners, and an external evaluation, involving external user groups participants of pilot users.

The project was presented and explained in many relevant events and is actively present in popular and technical social networks such as LinkedIn or Twitter.

Expected final results and use

SmartOpenData has discovered, transformed and published several biodiversity and environmental data sources. The datasets to be used were requested (if they were not open) and collected. In some cases, they were completely transformed before publication. Publishing those datasets facilitated full access to this useful information for SMEs, general citizenship, policy makers and other relevant stakeholders.

The analysis of Semantic Technologies allowed the project to understand and disseminate the Linked Open Data -LOD- Technologies inside the geospatial community. Besides this, the dialogue has been enriching in both directions. Indeed, the general principles of geographic information systems and geospatial data were also explained to technical partners, that now know better the natural, environmental and biodiversity contexts.

Contribution to standards has been one of the main results of the project and due to the creation of W3C-OGC working group, the SmartOpenData project has been active in this process.

Besides these technical results, the project has contributed to many other economic and administrative results. For example, the project can highlighted a greater transparency in public administrations, within our scope, through the improved visibility of environmental information. The external evaluation process carried out proved the interest of citizens and citizens' associations, SMEs and others relevant stakeholders regarding the published datasets and the applications developed using those technical tools enable them to contribute to environmental governance processes in the domains of transparency, knowledge management, accountability and responsiveness.

As an intangible, but no less important, result, SmartOpenData has achieved the development of a European common approach to LOD environmental data. In the first instance, this common approach was confined within the project consortium but it has served, in many countries and between several bodies and organizations, as the starting point for working groups that will continue the strategies and methodologies developed by SmartOpenData. Indeed, cooperation with other national and international research initiatives has been one of the main results of the project, for example, the project developed strong links with the GEOSS community due to participations in the AIP calls.

All of the planned technical, administrative and economic results of the project have been satisfactorily achieved. The final general results and impacts can be summarized in the following list:

Data shared by public bodies and final users: This is a major result as far as the level of data openness across European countries is very uneven. As a final result, a large amount of top quality shared linked open data has been published. SmartOpenData data providers have guaranteed an optimal origin to build the publication architecture as a starting point from which its Governance Model has served as a mechanism to attract new public bodies allowing new participations.

User engagement: User engagement is a key factor to get real impact. In order to promote it, SmartOpenData set up user groups in each area of interest and geographical location that provided very useful feedback to complete the final stages of the project.

Alignment with European protection environment trends and standards: The SmartOpenData data model and infrastructure were built upon using existing standards regarding geographic,

spatial, LOD and semantic standards. But in addition, SmartOpenData participated in this process through the W3C-OGC work group, promoted by this project, and others.

The project achieved the **Creation of an open source ecosystem** that through the consequent dissemination strategy and liaison plan has been offered to third parties and their customers to participate and benefit from the SmartOpenData outcomes.

Address of the project public website – www.smartopendata.eu.