This project is funded by the European Union

**Project Details**

**Project reference**:
620400

**Funded under**:
CIP

**Duration**:
36 months
01/10/2014 - 30/09/2017

www.energic-od.eu

# ENERGIC OD

# European NEtwork for Redistributing Geospatial Information to user Communities - Open Data

## D3.2: REPORT ON SURVEY AMONG PROJECT PARTNERS AND ON SURVEY WITH STAKEHOLDERS

This project is partially funded under the ICT Policy Support Programme (ICT PSP) as part of the Competitiveness and Innovation Framework Programme by the European Community

# Document Information

| | |
|---|---|
| **Project title** | European NEtwork for Redistributing Geospatial Information to user Communities - Open Data |
| **Call identifier** | CIP-ICT-PSP-2013-7 |
| **Project acronym** | ENERGIC OD |
| **Starting date** | 01.10.2014 |
| **End date** | 30.09.2017 |
| **Funding scheme** | Competitiveness and Innovation Framework Programme (CIP) - ICT Policy Support Programme - Pilot B |
| **Contract no.** | 620400 |

| | |
|---|---|
| **Deliverable no.** | D3.2 |
| **Deliverable name** | Report on survey among project partners and on survey with stakeholders |
| **Work Package** | 3 |
| **Nature**[1] | R |
| **Dissemination**[2] | PU |
| **Editor** | Nargess Kamali (GEOkomm) |
| **Authors** | Nargess Kamali (GEOkomm) |
| **Contributors** | ENERGIC-OD Consortium |
| **Date** | 29-05-2015 |

[1] R = Report, P = Prototype, D = Demonstrator, O = Other

[2] PU = Public, PP = Restricted to other program participants (including the Commission Services), RE = Restricted to a group specified by the consortium (including the Commission Services), CO = Confidential, only for members of the consortium (including the Commission Services).

# ENERGIC OD Consortium

National Research Council of Italy (CNR-IIA)
Italy

AED-SICAD Aktiengesellschaft (AED-SICAD)
Germany

Association Francaise pour l'Information Geographique (AFIGEO)
France

Bureau De Recherches Geologiques et Minieres (BRGM)
France

Global Infotech (NOW)
Italy

Centre National de la Recherche Scientifique (CNRS)
France

Depth France Sas (DEPTH)
France

Verband der GeoInformationswirtschaft Berlin/Brandenburg (GEOkomm)
Germany

Instytut Geodezji i Kartografii (IGIK)
Poland

Luftbild Umwelt Planung GmbH (LUP)
Germany

Politecnico di Milano (POLIMI)
Italy

Gesellschaft für Stadt-und Regionalplanung mbH (SRP)
Germany

Regione del Veneto (VEN)
Italy

Universidad de Zaragoza (UNIZAR)
Spain

Alkante Sas – Geonet (ALKANTE)
France

# Revision History

| Version | Date | Modified by | Comments |
|---|---|---|---|
| 0.1 | 29.05.15 | Nargess Kamali | First Issued |
| 0.2 | 05.06.15 | Michael Bauer | First revision and editing |
| 0.3 | 23.06.15 | Paolo Mazzetti | Revision and proofreading |
| 0.4 | 26.06.15 | Miguel Latre Abadia | Revision and proofreading |
| 0.5 | 08.07.15 | Mariella Liberti | Change to the new template |

# Table of Contents

# Index of Figures



# Index of Tables

# Abbreviations and acronyms

| Abbreviation / Acronym | Description |
|---|---|
| CA | Consortium Agreement |
| CIP | Competitiveness and Innovation Framework Programme |
| GIS | Geographic Information System |
| SDI | Spatial Data Infrastructure |
| VH | Virtual Hubs |
| DoW | Description of Work |
| EB | Executive Board |
| EC | European Commission |
| ENERGIC OD | European NEtwork for Redistributing Geospatial Information to user Communities- Open Data |
| GA | General Assembly |
| GrAg | Grant Agreement |
| ICT | Information and Communications Technologies |
| NEF | Negotiation Form Facility |
| PC | Project Coordinator |
| PO | Project Officer |
| PSP | Policy Support Programme |
| QC | Quality Control |
| R&D | Research and Development |
| WP | Work Package |
| WPL | Work Package Leader |
| GI | Geographic Information |

Table 1. Abbreviations and Acronyms

# EXECUTIVE SUMMARY

With numerous amounts of open data platforms and geographical information initiatives producing open data in various form and structure and under different licenses, it becomes increasingly exhausting to find the suitable data that is applicable and relevant to user's needs. In order to improve this condition an inventory of open data platforms and initiatives related to geographical information is necessary to provide information and metadata on best practices and concrete examples for availability of open GI. Therefore an exhaustive catalogue of open GI platforms and initiatives will be formed as a valuable source for the rest of ENERGIC-OD work packages.

This report explains in detail the structure of the survey and methodology that was defined in D3.1 [1] and evaluates each step of the survey. In the first step a list of open GI platforms is compiled by holding an interview with ENERGIC-OD partners and in the next step the major stakeholders were identified and interviewed subsequently. The result of this step is also compiled into the list of major open GI platforms. Furthermore, this list has been grown and expanded through internet research.

The next phase is to examine and assess the criteria, which were defined in D3.1, for all compiled platforms and to evaluate their efficiency and utilization in regard to open data platforms. In the course of this phase, some of these criteria were removed or adjusted for a better efficiency and suitability. Also during the evaluation process a demand for quality measurement for openness of platform emerged and as a consequence a quality measurement system for platform evaluation based on Tim Berners-Lee's 5-star scheme has been developed which aims towards improvement of the quality and accessibility of the open GI platforms.

# 1. Introduction

Work package 3 aims to provide an exhaustive catalogue of geospatial open data platforms, based upon the methodology and criteria defined for assessment of operational platforms on open GI in D3.1 [1]. In order to produce this catalogue, an inventory of open GI platforms deemed necessary. This inventory is compiled by conducting a survey in different steps, an initial survey with project partners was held in the first step, then through internet research as well as holding a survey with stakeholders more open GI platforms were acquired. Subsequently the criteria defined in D3.1 will be evaluated for each of these platforms in order to provide an exhaustive catalogue of best practices and concrete examples for availability of open GI. After this information has been processed and assessed, it will form a catalogue (the task for D3.3) that can facilitate comparison between the different platforms and initiatives. The catalogue will provide information and metadata on current and planned best practices, providing information for policy makers and all interested stakeholders. This catalogue will be accessible as an open service via ENERGIC-OD website.

The current document is a report on survey among project partners as well as with stakeholders in order to identify existing open GI platforms and initiatives.

## 2. Structure of the survey and Methodology

To be as comprehensive as possible, the survey is done in different stages by different means. This approach shown in Figure 1 has been developed and introduced in deliverable D3.1 with the intention of developing a method that provides an extensive list and can compile the open GI platforms by different means. In the following sections the approach depicted in Figure 1 is described in more detail.

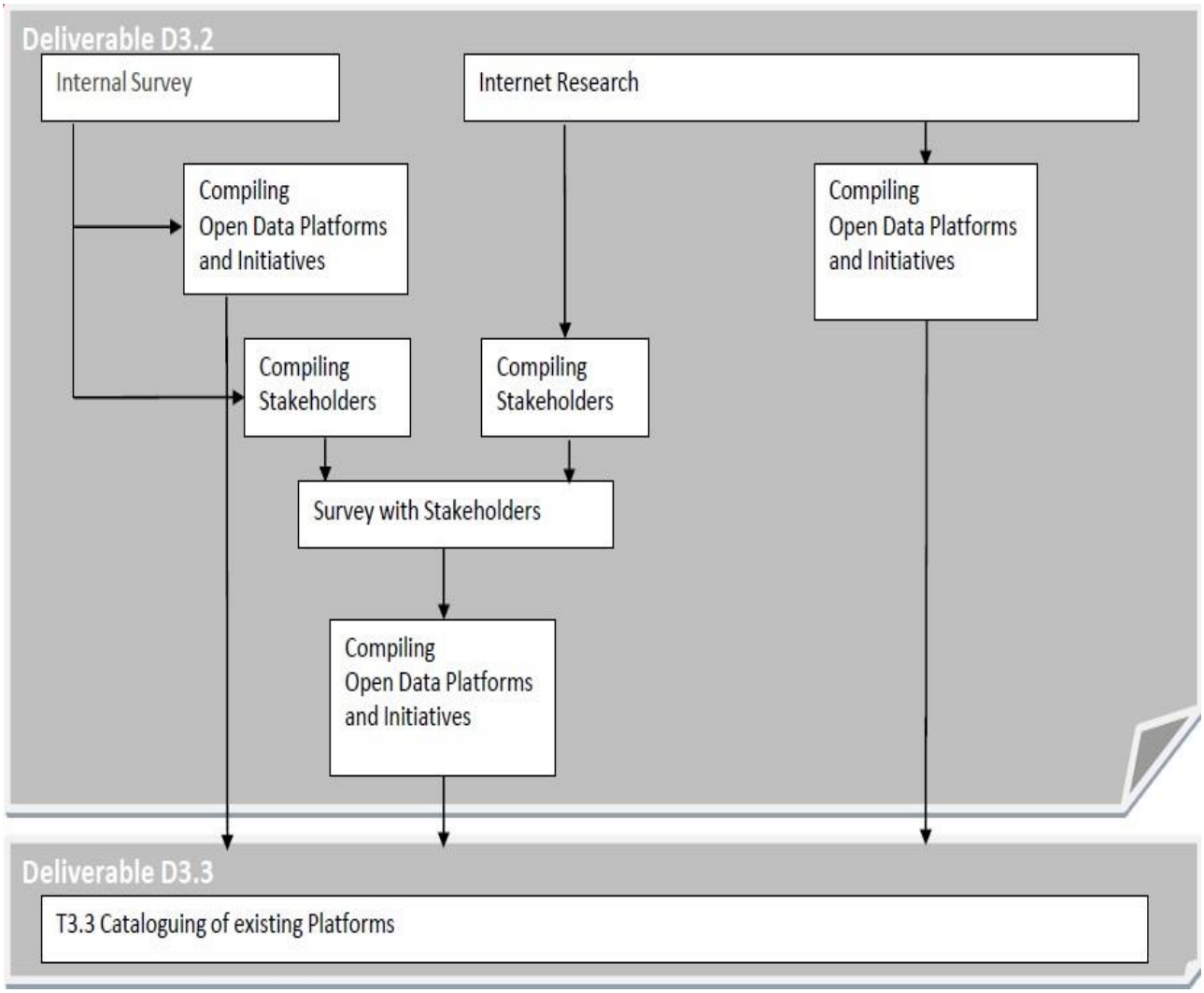


**Figure 1. Structure of the Survey**

### 2.1 Compiling Open Data Platforms and Initiatives

Internal survey was the first step in identifying platforms which provided open geo data. In this phase all the project partners were interviewed and asked to provide information concerning the notable open geo data platforms and initiatives in their respective countries as well as the major European platforms. An initial Excel form was prepared with 11 fields of criteria and was sent to the partners in order to fill out for the identified platform. The

outcome of each of these interviews were then compiled into the platform inventory. The result of this section was mainly an Excel file with arrays of open GI platforms listed within to be further evaluated with the criteria that have been defined in D3.1.

## 2.2 Compiling Stakeholders

Subsequently the project partners were asked to identify the suitable stakeholders for further expansion of the inventory of open GI platforms. Together with work package 7 a comprehensive list of stakeholders was compiled from surveying with project partners. Moreover the stakeholders relevant for work package 3 were identified for further inquiry by filtering the preliminary result of stakeholders' identification.

## 2.3 Survey with Stakeholders

The process of the stakeholders' identification from previous section was followed by interviewing the relevant stakeholders and collecting the major open GI platform that they could identify.

The result acquired in this step was added into the overall platform inventory (similar to 3.1).

## 2.4 Internet Research

The entire compilation process was repeated in Internet Research step. Here the search engines and governmental portals were used to discover some more GI open portals and finally the result of this part was similarly compiled and added into the open data platform list.

# 3. Assessing the criteria

The open data platform survey was followed by criteria assessment - at this point the criteria, which were developed for D3.1, were used to evaluate each of these platforms. In the course of this process, it became obvious that some of these criteria are not practical for the purpose of this project, some are redundant and some other are, when not impossible, remarkably hard and time-consuming to find and evaluate during the course of this work package. Consequently these criteria were modified and adjusted to some more functional measures. Also, where required, the criterion type has been adjusted in order to obtain an appropriate value. The modified criteria are shown below in Table 2 and an example of the differences and adjustments made to achieve Table 2 is shown in Table 3.

Furthermore tables "SERVICES" and "DATASETS" were merged into the "PLATFORM" table for higher efficiency. However, some of the criteria which were imported from the "Dataset" table are not really feasible when a platform is being assessed. The reason being that a platform incorporates many datasets which can be produced by different providers and therefore each dataset could hold different formats or origin, etc. in order to properly compile all these criteria a parser is required to go through all these datasets and fetch the appropriate information. The development and utilization of such parser could be

a task for the later stages of ENERGIC-OD project.

The table below shows the modified criteria for platform evaluation which was realized during the platform assessment.

| **PLATFORM** | | |
|---|---|---|
| ***ID*** | ***Description*** | ***Type*** |
| *Name* | Name or title of the platform or initiative | String |
| *Online Address* | Online address (URL) leading to the website of the platform or initiative | String |
| *Country* | Country of publisher of the platform or initiative<br>Intl. country codes (ISO 3166) | Varchar |
| *Publisher* | Publisher of the platform or initiative | String |
| *Scale* | Hierarchy level of the platform or initiative<br>*(Municipal, Regional, Federal/National, European, Global, Thematic)* | List |
| *Language* | Language or languages of the platform or initiative website<br>Intl. language codes (ISO 639) | Varchar |
| *Start Date* | The launch date (ISO 8601) of the platform or initiative | Date |
| *Ongoing* | Is the platform maintained and the data provided updated regularly? (ISO 8601) | Boolean[1]& Date |
| *Standard Compliant* | Standards and regulations applied by the platform<br>*(ISO, INSPIRE, etc.)* | List |
| *General Licence* | Does the platform use one licence for all datasets? | Boolean |
| *Registration* | Is a registration required in order to use all functionality of the platform | Boolean |
| *Search* | Search functionalities available to search for data<br>*(Metadata, Full text, Keywords, Topics)* | List |
| *Single access point* | Presence of a web portal through which all datasets of the platform can be accessed | Boolean |

[1] Boolean value can be defined here as 1, 0, and 99 as no information available on that!

| | | |
|---|---|---|
| *API* | Link to a developer interface, if provided | String |
| *Usability* | General assessment of the platform *(excellent, good, moderate, poor)* | String |
| *Transformation Service* | Coordinate transformation services offered by the platform? | Boolean |
| *Data Upload* | Upload possibility of new or altered data for the users | Boolean |
| *Openness* | Openness of platform (open, accessible and reusable) grading. (grade: 1-5) | Stars |
| *Geospatial Server* | Which geospatial service software is used by the platform? (Geoserver, MapServer, ArcGIS, Deegree) | List |
| *Metadata Server* | Which metadata service software is used by the platform? | String |
| *No. of Datasets* | Number of datasets included in the platform or initiative | Integer, Access Date |
| *No. of Users* | Number of unique users during the course of a month | Integer |
| *Service Address* | Machine to Machine (M2M) address of central data or metadata service | String |
| *Topic* | Thematic coverage of the platform *(Environmental, agricultural, traffic, etc.)* | List |
| *SRS* | Name and/or EPSG identification of spatial reference systems provided by the service | List |
| *Monitoring* | Are the services of the platform continually monitored concerning availability, usage and quality of service? | Boolean |
| *Geographical Extent* | Geographical area covered by the platform's datasets | String |
| *Contact* | The main contact for the platform | String |
| *Origin* | Source of data (public, private, crowd) | List |
| *Update Cycles* | Time interval (ISO 8601) data is updated | Timespan |
| *Metadata Schema* | Applied standards for geospatial metadata (Dublin Core, INSPIRE, etc.) | List |

| | | |
|---|---|---|
| *Structure* | Concept the data is structured by – especially relevant for crowd sourced data | String |
| *Type* | Type of datasets<br>(Coverage, vector data, etc.) | String |
| *Format* | Format of datasets<br>(GML, PNG, JPG, JSON, etc.) | List |

Table 2. Table of the criteria for platform assessment

| Criteria before evaluation and modification | Criteria after evaluation and modification |
|---|---|
| *Federal State* | Having both *Publisher* and the *Country* of publisher of the platform among the criteria list, this criterion seemed to be redundant |
| *Host* | Having both *Publisher* and the *Country* of publisher of the platform among the criteria list, this criterion seemed to be redundant |
| *OGC-Services* | This criterion is covered by *Standard Compliant* and is redundant here |
| *Version* | It is very time consuming to find the version of a service because most of the time the version is not mentioned on the platform's website |
| *Protocols* | This criterion is covered by *Standard Compliant* and is redundant here |

Table 3. An example of the modified criteria

## 3.1 The open data platform survey spreadsheet

In the following table an example of the OD platform survey together with their evaluated criteria are shown (Table 4). This information will hopefully be used later on to improve the quality of the platform on the one hand and to provide the right data for users on the other hand.

| | |
|---|---|
| Name | GEOSS |
| Scale | International |
| Online Address | http://www.geoportal.org/web/guest/geo_home_stp |
| Openness | ***** |
| Geospatial Server | ArcGIS |
| Standard Compliant | OGC web services, CSW, ISO, GSF(GEOSS Service |

| | |
|---|---|
| | Factory components), KML, GML |
| Metadata Server | ArcGIS Server Geoportal Extension |
| Language | English |
| Single Access Point | Yes |
| License | have their own license |

Table 4. Example of an OD platform survey

## 3.2 Problems

During the process of platform survey with partners and stakeholders and also throughout evaluation phase of the criteria for the platforms some complications occurred that are described distinctively in this section.

- **Poor Metadata**: the main challenge during the OD platform assessment was to find information relevant to the criteria which were hard to find mainly due to poor, non-existent or inaccessible metadata. This deficiency has led us to develop a quality measure for evaluation of geoportal and openness of a platform, based on Tim Berners-Lee 5-star scheme for open data.
- **License Ambiguity**: the licences of datasets were not clearly presented in some platforms and this fact makes for “ambiguousness of open data” and causes limitation for using the data in different applications due to the right of use as the metadata for each dataset has to be consulted to find out how the data is permitted to use.
- **Language**: what made the survey of the open data platforms challenging was the language of the web portals. Many of the platforms use solely the native language of the country or the region and therefore make it difficult to navigate through their portal.
- **Redundancy**: receiving redundant data from partners in the same country and having difficulties receiving any information from some other partners are challenges that were expected but nonetheless troublesome.
  And last but not least:
- **Stakeholders’ Survey**: these experiences naturally were multiplied during open data survey with external stakeholders, namely difficulties finding the stakeholders, contacting stakeholders who weren’t interested in the project and collecting redundant platforms in the same country.

# 4. Openness of data versus openness of platforms

## 4.1 Open data 5-star scheme comparison with Open Data Certificates

All above-mentioned drawbacks and the need for a platform qualification measurement has led to the development of a platform openness scheme - an adoption of both Tim Berners-Lee's 5-star scheme for open data and Open Data Certificates.

In order to provide a better insight for platform openness development and add the complementary background information, the Open Data Certificates is compared shortly with 5-star scheme in the following section.

The 5-star deployment scheme essentially measures how well data is integrated into the web.

> *"1-star" data is published in proprietary formats that users must download and process.*
>
> *"5-star" data can be accessed online, uses URIs to identify the resources in the data, and contains links to other sources. [2]*

The 5-star scheme is focused on how data is published and what formats and technologies are used for these data. Assessing whether a dataset is published at 2, 3 or 4 stars requires some insight into how the data has been published, which can be difficult for a non-technical person to assess and therefore is best used as a technical roadmap of the technical aspects of data publishing.

The Open Data Certificates measures how effectively a dataset is shared in terms of ease of reusability. This extends more than just the technical aspects and into integrating the right of use and licensing, documentation, and guarantees about availability.

> *A certificate therefore offers a more rounded assessment of the quality of publication of a dataset and acts as a reference sheet containing information of interest to reusers of a dataset. [2]*

Therefore assessment of a dataset gives the data publisher the opportunity to improve their publishing process efficiently. Furthermore the advantage of using this certificate is that the reusers can spare the time and effort that they might have spent otherwise on digging the publisher's website, and instead search in order to realize whether a dataset is suitable for their needs. Hence the certified datasets are easier to search and more likely to provide the useful data.

Nevertheless the 5-star scheme and the Open Data Certificate both intend to improve the quality and accessibility of published data by requiring open licenses and using standard formats for published data. And that is also what we tried to develop for the platform quality evaluation and openness measurement. [3]

## 4.2 Openness of Platform

In order to develop a quality measurement system for platform evaluation, the 5-star scheme introduced by Tim Berners-Lee was analyzed in more depth and tailored to fit the requirements by which a platform could be evaluated. The proposed 5-star scheme for openness of platform is shown in Table 5.

| Tim Berners-Lee 5-star Scheme | Openness of Platform Adaptation | Stars |
|---|---|---|
| 1-star means *available* on the Web. | For a platform, it means that a user could search, evaluate and access datasets through a Graphical User Interface but there is no service interface available. | The platform receives 1 star |
| 2-star means *structured*. | For a platform it means that discovery and access are available through web services for machine-to-machine communication. This might be proprietary or legacy services. | The platform receives 2 stars |
| 3-star means *non-proprietary.* | For a platform it means that discovery and access are available through at least one standard service interface. | The platform receives 3 stars |
| 4-star means *referable.* | This condition doesn't apply on a platform, instead it is used to measure the usability of the platforms by having a consolidate license. Furthermore, harmonization could be achieved at this level, which means all datasets should be produced according to the same specification and same CRS, etc. | The platform receives 4 stars |
| 5-star means *linked.* | Similar to the previous condition this doesn't hold true for platform openness measurement. Therefore to achieve a higher performance and usability from a platform, multilingualism is included as a semantic service for discovery in order to achieve 5 stars | The platform receives 5 stars |

Table 5. Platform Openness

Openness of platform aims to improve the quality and accessibility of the open datasets that are provided by the platform. Additionally a virtual hub could connect 2-4 star platforms and from this congregation a 5-star Meta platform could be created.

# 5. References

1. D3.1, ENERGIC-OD Consortium, 2014. Methodology and a list of criteria for assessment of open data platforms and initiatives across EU.
2. Comparing the 5-star scheme with Open Data Certificates

   https://theodi.org/blog/5-star-open-data-certificates-tim-berners-lee

3. Open Data Certificate

   https://certificates.theodi.org/about