

European Data Portal

Landscaping Method – Overview

Second edition 2016

The purpose of this document is to present an overview of the landscaping method adopted in the European Data Portal project (EDP). The intention of the EDP consortium is to have a document ready-at-hand to present the landscaping approach to Member States, partner projects and other stakeholders.

Table of Contents

1	Landscaping Method	2
1.1	Objective of the landscaping within EDP	2
1.2	Working Definitions	2
2	Approach for the Landscaping.....	3
2.1	Required input	3
2.1.1	Work Approach.....	3
2.1.2	Overview of Indicators	3
2.1.3	Additional Research Work.....	5
2.2	Expected Output	6
2.3	Scoring	7
2.3.1	Political situation	7
2.3.2	Presence of an Open Data policy	7
2.3.3	Use of Open Data	8
2.3.4	Impact of Open Data	8
2.3.5	Portal features.....	8
2.4	Timing of Work	9
	Annex I: Domains listed in the G8 Open Data Charter	10

1 Landscaping Method

1.1 Objective of the landscaping within EDP

The *European Data Portal project* (EDP) covers a full set of activities including upgrading of the <http://www.europeandataportal.eu/> portal infrastructure and the provision of a series of added value services to accompany the European countries. Our overarching approach to support European Countries as they mature on their Open Data journey is to offer a combination of common support material and training as well as targeted support for those that are still in the early stages of the journey.

In 2015, the first step in assessing where the countries stand on their journey to implement Open Data was completed. The results of this assessment are summarized in the report '[Open Data Maturity in Europe](#)'. This activity is commonly called landscaping and is repeated to compare the results of 2015 with the current status in 2016. Improvements, novel features and approaches can therefore be documented and compared over time, in a consistent manner. Landscaping involves assessing where different European countries stand with regard to their Open Data policies and, more importantly, their publishing practices. The results provide an indication of the progress made in one year time and will help to further establish where to provide common support activities and where to offer further specialised training and coaching assistance within a safe learning environment.

In order to serve the countries as they mature on their Open Data journey, it is crucial to collect information on their progress and challenges from the Member States themselves.

1.2 Working Definitions

This section provides a working definition of what is to be understood as Open Data:

Open (Government) Data refers to the information collected, produced or paid for by the public bodies (also referred to as Public Sector Information) and made freely available for re-use for any purpose. The license might also say:

- that people who use the data must credit whoever is publishing it (this is called attribution)
- that people who mix the data with other data have to release the results as Open Data as well (this is called share-alike)

These principles for Open Data are described in detail in the Open Definition.¹

Public Sector Information is information collected by the public sector. The Directive on the re-use of Public Sector Information² provides a common legal framework for a European market for government-held data (Public Sector Information).

The specific scope of Open Data used within the European Data Portal project is data published by public administrations or on their behalf. The focus is not on community data.

The most common definition of a dataset: it corresponds to the contents of a single database table, or single statistical data matrix. Thus, referring to an organised collection of data that can be accessed on a data portal.

¹ <http://opendefinition.org/>

² Directive 2003/98/EC <http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2003:345:0090:0096:EN:PDF>

2 Approach for the Landscaping

2.1 Required input

First, the input to assess the Open Data Maturity of countries in Europe will be described. The work approach, description of the indicators and additional research work will be discussed.

2.1.1 Work Approach

Each year, the EDP Consortium will conduct a maturity assessment to identify where European countries stand on their Open Data journey at a fixed point in time. The first assessment was delivered in October 2015. Open Data maturity is measured against two key indicators. The first indicator called “Open Data Readiness” comprises presence of specific Open Data policies, licensing norms, the extent of coordination at national level, when it comes to guidelines and setting common approaches, as well as whether the revised PSI directive has been transposed. A second scoring consists of another composite indicator measuring “data availability.” It is made of usability of the portal with regard to availability of functionalities, the use made of the data available, overall re-usability of data as well as spread of data across high value domains. These indicators are depicted below in Table 1.

Open Data Maturity Assessment							
Open Data Readiness					Portal maturity		
1. Presence of Open Data Policy	2. Licensing Norms	3. Extent of coordination at national level	4. Use of Data	5. Impact of Open Data	6. Usability of the portal	7. Re-usability of data	8. Spread of data across domains

Table 1: Open Data Maturity Indicators

For each maturity assessment, a set of complementary sources will be used. Indeed, assessing maturity implies seeking both quantitative and qualitative information. As far as possible, the consortium will re-use existing information already collected by for instance:

- The Open Data Barometer
- The Open Data Index
- The metadata harvested by the European Data Portal itself
- In addition, desk research will be conducted.

However, most of these sources do not offer a broad coverage of the policy aspects related to the implementation of Open Data. It is necessary to collect further qualitative information to assess the level of Open Data readiness of the different European countries.

To meet the goal of the landscaping and enable tailored support to be brought to the countries, the members of the PSI expert group are kindly invited to fill in the questionnaire within this document. It is structured into eight chapters: Information on the respondent, presence of Open Data policy, the use of Open Data, the impact of Open Data, best practices, challenges and barriers, portal features and open additions.

The output of the Landscaping is presented in section 2.2.

2.1.2 Overview of Indicators

This section presents the detailed indicators below. Open Data Readiness is mostly gathered via the input of the questionnaire. Should countries not wish to fill this section in, they will be kindly invited to confirm that the responses delivered in 2015 are the ones to be used for the 2016 landscaping. We

believe a lot has happened since Spring 2015 and encourage the countries to provide us with their responses in order to benefit from the most accurate as possible assessment.

Open Data Readiness

Indicator	1. Presence of specific Open Data policy
1.1	Open Data policy and policies supporting re-use are in place. Open Data policy is different from the PSI policy. A national strategy on Open Data does exist.
1.2	National, but also regional/local portals are present. The data holders are able to upload the data themselves, the frequency of collection is provided and there is a pre-defined approach to ensure data is up-to-date.
1.3	Priority domains are identified. The public administration is using data themselves for decision-making, but promoting the use by others as well by organizing events.

Indicator	2. Licensing norms
2.1	Data is available free of charge.
2.2	Data is open licensed.
2.3	A national data policy provides for a standard license (or suite of licenses) that public sector bodies are encouraged to avail themselves of when allowing PSI re-use.

Indicator	3. Extent of coordination at national level
3.1	National guidelines on the publications of PSI are in place.
3.2	Numerous regions and/or cities run their own Open Data initiatives, like portals or specific policies, and are integrated on the national portal. NGO's and private companies are allowed to share their datasets on the portal.

Indicator	4. Use of the data
4.1	Overview of portal traffic statistics: number of unique visitors, visitor profile, percentage human, percentage foreign, and API access.

Indicator	5. Impact
5.1 Political Impact	
5.1.1	High impact on government efficiency and effectiveness.
5.1.2	High impact on transparency and accountability in the country.
5.2 Social impact	
5.2.1	High impact on environmental sustainability in the country.
5.2.2	High impact on increasing the inclusion of marginalised groups in policy making and accessing government services.
5.3 Economic impact	
5.3.1	Multiple macro-economic studies assessing the market value of Open Data are done as well as studies regarding better service delivery or looking at related subjects.
5.3.2	The funding model is known.

Portal Maturity

The first landscaping assessment did not include questions regarding portal maturity. Those indicators were obtained by going through the data portals. This second round offers the countries the opportunity to provide this information themselves, if they wish to. The research team will equally go through the portals. The portal maturity indicators are shown below.

Indicator	6. Usability of the portal
6.1	Feedback mechanisms are available on the portal to comment on data set quality and get a discussion going.
6.2	Users can access data sets, search, download and contribute themselves.
6.3	A distinction is made between human and API access, indicating that an API is available.

Indicator	7. Re-usability of the data
7.1	The proportion of data that is machine-readable is known (Open Data Monitor).
7.2	All data is available in one – bulk – download
7.3	File formats are searchable and it is known which file format is available the most.

Indicator	8. Spread of data across domains
8.1.	Data sets are numerous and up-to-date
8.2	Multiple organisations provide a lot of data sets
8.3	Data sets are searchable by domain with many different domains present

2.1.3 Additional information

Additional information will be conducted and included in the landscaping questionnaire addressed to the PSI Expert Group. This information will be used to complement the information for the country fiches. This information is not scored. The material collected will support in assessing the political, social and economic impact as well as best practices and barriers. Finally, the research team will also leverage the eGovernment context to obtain an overall picture of the digital landscape of the countries assessed.

Indicator	A. Best practices and examples
A.1	Best practices and examples

Indicator	B. Barriers
B.1	What are the barriers to the further publication and use of Open Data in this country?

Indicator	C. Use of the Portal
C.1	Overview of portal traffic statistics: number of unique visitors, geographic coverage
C.2	Top ten most accessed datasets.
C.3	Top ten stakeholders accessing the portal

Thus, to meet the goal of the second landscaping and demonstrate the progress achieved by the countries, the PSI expert group contacts are kindly invited to fill in the questionnaire again. It is structured into seven chapters: Information on the respondent, presence of Open Data policy, the use of Open Data, the impact of Open Data, best practices, challenges and barriers, open additions.

2.2 Expected Output

The output of the landscaping will consist of Insight report n2 presenting where each and every country stands on their Open Data Journey compared to 2015. The level of maturity will determine the level of support the country should be given and how this is articulated with our offering. This will support the development of a structured and collaborative journey with the countries in supporting them mature in their Open Data Journey:

Beginners: the initial steps have been made, but they are still struggling with basics around availability, accessibility, portal functionalities remain limited and there is a limited coverage in terms of datasets.

Followers: the basics are set, including a clear vision and there are advanced features on the portal. However, the approach to the release of data is very much in silo and remains limited.

Leaders – Trend Setters: these are the most advanced and have solid portals with elaborate functionalities and coordination mechanisms across domains.

In addition, the 2015 country fiches depicting the specific situation of each country will be updated with the new information. Country fiches also contain further information with regard to the impact, best practices and main barriers faced for further publication of data and its re-use. An example of the Spanish 2015 country fiche is shown below.

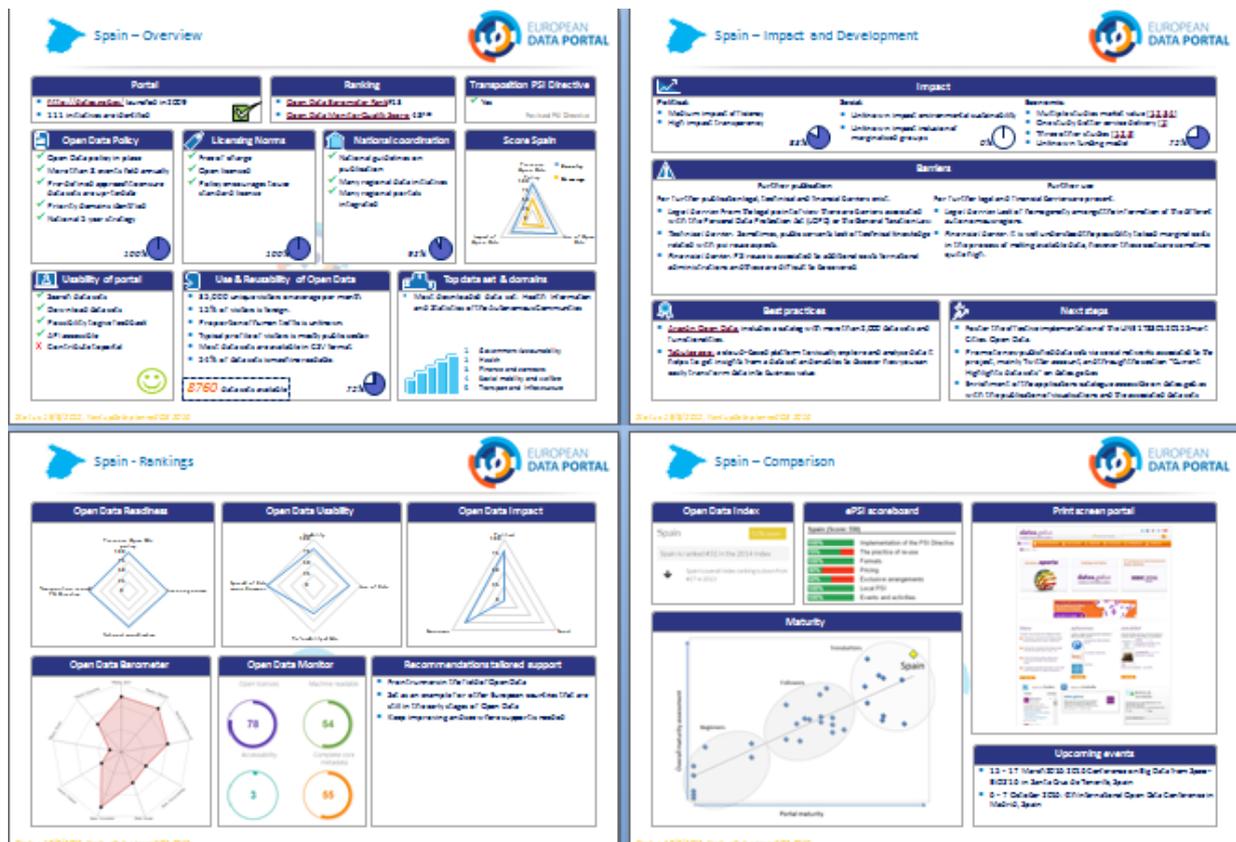


Figure 1: Example of the country fiche of Spain in 2015

The country fiches are updated on a yearly basis, thus this assessment contributing to the second edition. The results will be published in the form of a dashboard and presented on the European Data Portal. Country fiches will also be made publicly available. To guarantee they are fit for publication, a validation step will be conducted with the individual countries from mid-August to mid-September.

2.3 Scoring

A short explanation of the scoring mechanism used to score Open Data Maturity in Europe is given. In total, countries are able to score a maximum of 1250 points. The most important part in terms of scoring includes the political situation. The Open Data policy, use and impact add up to a maximum of 1000 points. The portal features account for a maximum of 250 points. The political part weights more compared to the technical implementation of the portal, as governmental support is a very important precondition for a successful Open Data implementation. A result of the decision to score the national portal is that countries without a national portal are not able to score more than 80% of the maximum number of points.

2.3.1 Political situation

The maximum of 1000 points is distributed with 500 points for the presence of an Open Data policy and 500 points for the use and measuring the impact of Open Data. The presence of a legal foundation and the availability of guidelines are essential requirements for the development of Open Data, and therefore scored equally compared to use and impact combined. Going in more detail of the scoring of the presence of an Open Data policy, 300 points are assigned to the Open Data policy itself and 200 points to the licensing norms and extent of coordination at national level.

2.3.2 Presence of an Open Data policy

Assessing the Open Data Policy is done via a series of 12 questions with a maximum score between 10 and 50 depending on the importance of the questions. For example, countries are able to score 50 points when their data collection frequency is daily. Re-users of Open Data are highly depending on up to date or even real-time data, which makes the frequency of data collection an important condition for re-users to use the data. Another important question with a maximum score of 50 points is the number of events held annually. Events are among the best ways to raise awareness around the existence of Open Data. The countries that hold more than eight events annually score 50 points. Compared to 2015, three additional questions have been added that can be scored with a maximum of 30 points. These questions enable a deeper understanding of the promotion made of the re-use of Open Data. The number of events organized has been adjusted based on whether the countries are considered of small size (population < 9 million); average (population between 9 and 35 million) and large (population > 35 million). In addition two questions are added addressing changes in the Open Data policy and approach to Open Data since 2015.

The licensing norms section only consists of three questions with a maximum score of 70 points. Data that is available free of charge stimulates re-use the most and is scored with 30 points. An open licence is important to be able to use the data as Open Data with a maximum score of 25 points. The availability of a national policy providing a standard license is scored a bit lower, as it does not prevent the publications or re-use of data.

Extent of coordination at national level results in a maximum of 130 points. The five questions are scored between 10 and 40 points. The most points can be achieved by having many regional events. Many local initiatives in a country indicate the spread of Open Data across the country. When a country has integrated all regional portals in their national portal, they score another 60 points. One additional

question is added addressing the possibility for NGO's and private companies to share their data on the national Open Data portal.

2.3.3 Use of Open Data

Five questions are asked regarding the use of Open Data. A country scores 80 points when the number of visitors on their portal per month is above 0.05% of the number of inhabitants the country has. This scoring method thus changed compared to the assessment in 2015. Such an approach mitigates any discrimination based on the size of the country. This question is linked to the maturity of the national Open Data portal, but is more focused on the re-use of Open Data than the portal itself. An API stimulates re-use by developers and adds another 40 points to the total score. The other three questions have a comparable weight. Compared to 2015, four questions are added to this section. Those questions can be scored with an additional 60 points. The questions ask about changes in re-use since 2015, the launch of monitoring activities, support of re-use and further communication activities promoting Open Data.

2.3.4 Impact of Open Data

Impact of Open Data is divided between political, social and economic impact. The questions address whether these impacts have been measured as well as how they are perceived. The weight of political and economic impact is twice the weight of social impact, as the social impact is difficult to measure and in general more perceived as an indirect benefit of Open Data. However, all questions of the different impact types have a similar weight. One question was changed compared to 2015. As there were two questions around the impact of increased government efficiency and effectiveness, the first questions changed to the launch of activities to monitor the impact of Open Data. The total maximum scored is the same as for 2015.

2.3.5 Portal features

In 2015, the portal features were assessed by the research team. This year, countries have the opportunity to provide further information on a voluntary basis. There have been slight changes in order to refine the estimation of maturity of the national Open Data Portal. Usability of the portal and re-usability of the portal had an equal maximum number of 100 points, but in 2016 those points are changed to 60 and 140 points. Both topics determine the user-friendliness of the portal and stimulate re-use. However, the re-usability is more important for re-users as it determines whether they will return to the portal to easily find the data they need. The spread of data across domains is important to attract visitors with different backgrounds, but less important to keep them on the portal. Hence, the maximum score of 50 points for the spread of data across domains. The most important question in this part of the scoring is the proportion of files that is available in machine-readable format with a maximum of 60 points when 90% of the files or more is machine-readable. Developers of web or mobile applications need files that they can automatically process as input for their product or service. All other questions are weighted equally. The research team will crawl the different portals in an automated fashion in order to gather this information. As can be seen in the above, the scoring method does not discriminate smaller from larger countries as quantity is not measured, rather availability of features.

2.4 Timing of Work

The picture below depicts the overall timing of the completion of the landscaping.

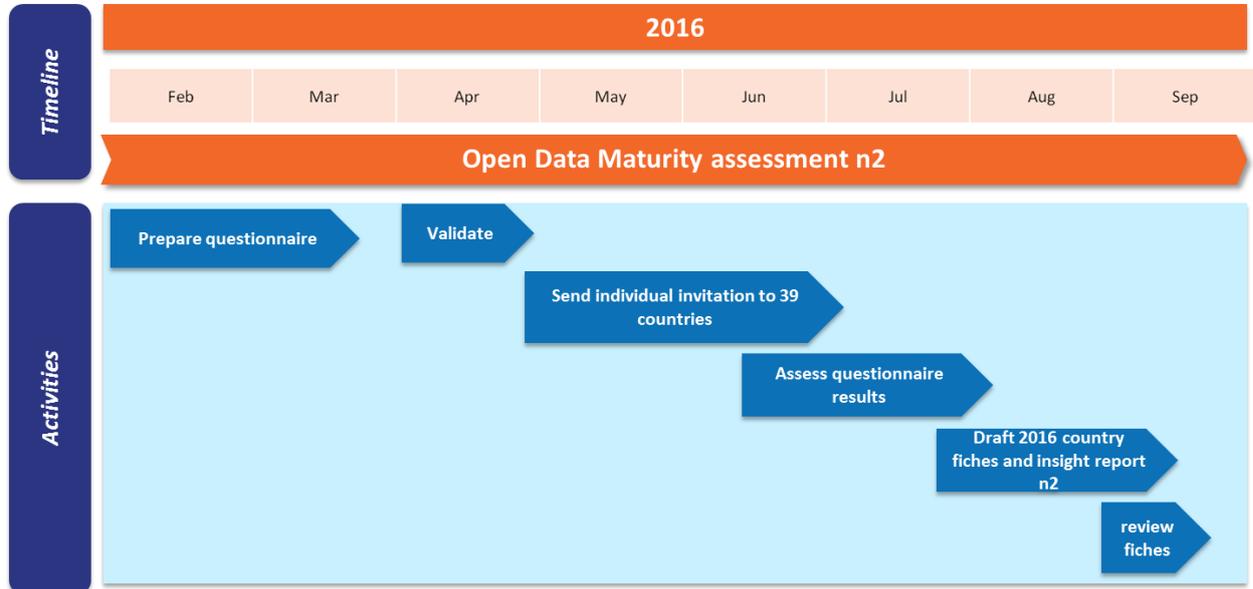


Figure 2: Timing of Work

Annex I: Domains listed in the G8 Open Data Charter

To adopt an internationally recognised approach in structuring datasets, the consortium uses the domains listed in the G8 Open Data Charter³.

Data Category* (alphabetical order)	Example datasets
Companies	Company/business register
Crime and justice	Crime statistics, safety
Earth observation	Meteorological/weather, agriculture, forestry, fishing, and hunting
Education	List of schools; performance of schools, digital skills
Energy and environment	Pollution levels, energy consumption
Finance and contracts	Transaction spend, contracts let, call for tender, future tenders, local budget, national budget (planned and spent)
Geospatial	Topography, postcodes, national maps, local maps
Global development	Aid, food security, extractives, land
Government accountability and democracy	Government contact points, election results, legislation and statutes, salaries (pay scales), hospitality/gifts
Health	Prescription data, performance data
Science and research	Genome data, research and educational activity, experiment results
Statistics	National Statistics, Census, infrastructure, wealth, skills
Social mobility and welfare	Housing, health insurance and unemployment benefits
Transport and infrastructure	Public transport timetables, access points broadband penetration

³ https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/207772/Open_Data_Charter.pdf