

Pan-European Open Data Portal

Landscaping Method – Overview

31 March 2015

The purpose of this document is to present an overview of the landscaping method adopted in the Pan-European Open Data Portal Project (ODPP). The intention of the ODPP consortium is to have ready-at-hand a document 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 ODPP	2
1.2	Working Definitions	2
2	Approach for the Landscaping	2
2.1.1	Work Approach.....	2
2.1.2	Overview of Indicators.....	4
2.1.3	Additional Research Work	6
2.2	Expected Output	7
2.3	Timing of Work.....	8
	Annex I : Domains listed in the G8 Open Data Charter	9

1 Landscaping Method

1.1 Objective of the landscaping within ODP

The *Open Data Portal Project* (ODPP) covers a full set of activities with the ambition of deploying a pan-European data portal infrastructure including the provision of a series of added value services to accompany the European countries.

Our overarching approach to supporting 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.

A first step therefore consists in assessing where the countries stand on their journey to implement open data, increase the publication of data sets and support the re-use of the data made available.

This activity is commonly called landscaping. This involves assessing where different European countries stand with regard to their open data policies and, more importantly, their publishing practices. This will serve as the knowledge base to help to build our leaders' network, and to establish where to provide common support activities and where to offer further specialised training and coaching assistance.

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 also release the results as open data (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 EU Open Data Portal project is data published by public administrations or on their behalf. The focus is not on community data.

2 Approach for the Landscaping

2.1.1 Work Approach

Each year, the ODP Consortium will conduct a maturity assessment to identify where European countries stand on their open data journey at a fixed point in time. Open data maturity will be measured against two key indicators. The first composite indicator called “open data readiness” will

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>

comprise 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 in 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:

Open Data Maturity Assessment							
Open Data Readiness				Data availability			
1. Presence of Open Data Policy	2. Licensing norms	3. Extent of coordination at national level	4. Transposition of the revised PSI directive	5. Usability of the portal	6. Use of Data	7. Re-usability of data	8. Spread of data across domains

Figure 1 Open Data Maturity Indicators

The different steps required for the realization of the Landscaping are described below:

a. Step 1: Leveraging existing open data monitoring activities

The main approach is to focus on the re-use of the following existing and ongoing landscaping activities

- The Open Data Monitor³
- The Open Data Barometer⁴

We also propose the re-use of several of indicators used by ePSI⁵. While the ePSI activities end in 2015, Capgemini will conduct the research itself for the specifics of sub-indicator 4 under readiness of the country.

b. Step 2: Complementing existing monitoring with desk research

To complement the data from the existing studies, a questionnaire will be created. The respondents will consist in the individuals from the PSI expert group⁶ which is chaired by the European Commission. In addition, the work done by the Share PSI 2.0 group will also be leveraged as they are working on the collection of best practices and barriers.

c. Step 3: Producing a detailed overview of the findings

d. Step 4: Visualising the results

The expected output of the Landscaping is presented in section 1.4.

Acronyms used:

ODI: Open Data Index

ODB: Open Data Barometer

ODM: Open Data Monitor

3 <http://project.opendatamonitor.eu/>.

4 <http://www.opendatabarometer.org/>.

5 <http://www.epsiplatform.eu/>. For more information, see chapter 6

6 <http://ec.europa.eu/digital-agenda/en/news/public-sector-information-group-main-page>

2.1.2 Overview of Indicators

This section presents the detailed indicators below.

1. Indicators: Readiness of the country

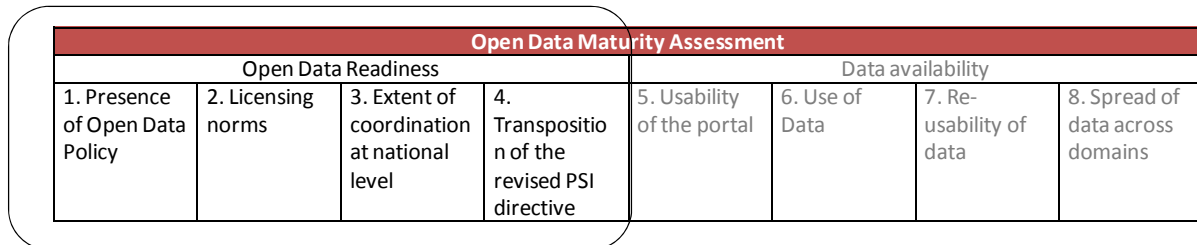


Figure 2 The first four composite indicators for the readiness of the country

Indicator	1. Presence of specific Open Data policy	Method
1.1	Do the Member States have a specific policy on Open Data? Are there policies supporting the re-use of public sector information?	PSI Questionnaire, ePSI, desk research
1.2	To what extent are city or regional governments running their own open data initiatives?	PSI Questionnaire, ePSI, desk research
1.3	To what extent is government directly supporting a culture of innovation with open data through competitions, grants or other support? Have there been at least 4/8/12 national or inter-regional events held annually to promote Open Data and PSI re-use (organized by whichever organization or group)? (e.g. hackathon, events)	PSI Questionnaire, ePSI, desk research

Indicator	2. Licensing norms	Method
2.1	Is the data available for free?	ODI & ODB
2.2	Is the data open licensed? (Proportion of distribution with open license = ODM)	ODI, ODB, ODM
2.3	Is there a national data policy providing for a standard license (or suite of licenses) that public sector bodies are encouraged to avail themselves of when allowing PSI re-use? If yes than what is it?	PSI Questionnaire, ePSI, desk research

Indicator	3. Extent of coordination at national level	Method
3.1	Is there a national portal? Is the portal overseen by the government?	PSI Questionnaire, ePSI, desk research
3.2	Governance: How is the collection, use and accessibility of Open Data governed?	PSI Questionnaire, ePSI, desk research

Indicator	4. Transposition of the revised PSI directive	Method
4.1	Has the revised PSI directive been transposed into national	ePSI, Capgemini research

law by the deadline of 18 July 2015?

2. Indicators: Data availability

Open Data Maturity Assessment							
Open Data Readiness				Data availability			
1. Presence of Open Data Policy	2. Licensing norms	3. Extent of coordination at national level	4. Transposition of the revised PSI directive	5. Usability of the portal	6. Use of Data	7. Re-usability of data	8. Spread of data across domains

Figure 3 The three composite indicators for the readiness of the country

Indicator	5. Usability of the portal	Method
5.1	Are feedback mechanisms available to the user to give his opinion on the service? (any type: user satisfaction monitoring, polls, surveys, etc).	Capgemini research
5.2	Users can both download and contribute the datasets	Capgemini research
5.3	Users can only read and download datasets	Capgemini research
5.4	Users can only read or access the content on the portal	Capgemini research
5.5	Is there a distinction made between human and machine accessibility of the Portal? Can machines access your Portal via a specific API ⁷ ?	Questionnaire PSI group

Indicator	6. Use of the data	Method
6.1	Overview of portal traffic statistics: number of unique visitors, geographic coverage	Questionnaire PSI group
6.2	Top ten most accessed data sets	Questionnaire PSI group
6.3	Top ten stakeholders accessing the portal	Questionnaire PSI group
6.4	What proportion of traffic is generated by a human and what is accessed by the API (i.e. Machine traffic).	Questionnaire PSI group

Indicator	7. Re-usability of the data	Method
7.1	Is the data available in machine readable format? What proportion?	ODM
7.2	Is the data available in bulk? ⁸	ODM
7.3	Frequency of distributions by file format Values: [Compressed/Archive (zip, gz, tar, rar), HTML, PDF, Word, Excel, CSV, TSV, JSON, XML, RDF (all serialisations), other or unknown Notes: based on catalogue metadata	ODM

⁷ Application Programming Interface

⁸ The data should be available as a complete set. The entire register should be available for download. A web API or similar service is not a substitutes for bulk access.<http://opendatahandbook.org/en/how-to-open-up-data/make-data-available.html>

Indicator	8. Spread of data across domains	Method
8.1 Spread of data per domain		
8.1.1	EU measures over the aggregate	ODM
8.1.2	Data volume per country	ODM
8.1.3	Is the dataset up to date? (<i>is this useful to measure?</i>)	ODB, ODI
8.1.4	What data is targeted? What data are we talking about – different types of domains?	Capgemini questionnaire, Machine analysis ODM
8.2 Spread of data across stakeholders and level of government		
8.2.1	Proportion of catalogues by sector of publishing organization <ul style="list-style-type: none"> • Government/public sector • Other non-commercial/3rd sector • Commercial • Mixed 	ODM
8.2.2	Are there at least two / six / twelve local or regional governments that publish at least 10 PSI data sets for re-use,	ePSI

2.1.3 Additional Research Work

Additional research work 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 that can be used for the peer learning. 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. Impact	Method
A.1 Political Impact		
A.1.1	To what extent has open data had a noticeable impact on increasing government efficiency and effectiveness?	ODB, PSI Questionnaire, desk research
A.1.2.	To what extent has open data had a noticeable impact on increasing transparency and accountability in the country?	ODB, PSI Questionnaire, desk research
A.2 Social impact		
A.2.1	To what extent has open data had a noticeable impact on environmental sustainability in the country?	ODB, PSI Questionnaire, desk research
A.2.2	To what extent has open data had a noticeable impact on increasing the inclusion of marginalised groups in policy making and accessing government services?	ODB, PSI Questionnaire, desk research
A.3 Economic impact		
	Have there been macro-economic studies assessing the market value of data? (e.g untapped business opportunities represented by the use of open data? Or information on better service delivery for users of public services)	ODB, PSI Questionnaire, desk research
9.5	Is open data used in decision making? How?	PSI Questionnaire, desk research

Indicator	B. Best practices and examples	Method
B.1	Best practices and examples	Findings from existing studies (re-use), PSI Questionnaire, desk research
Indicator	C. Barriers	Method
C.1	What are the barriers to the further publication and use of open data in this country?	Findings from existing studies (re-use), PSI Questionnaire, desk research
Indicator	D. Overall eGovernment context	Method
C.1	What is the overall eGovernment context? How can European governments improve to benefit from innovative public services?	DG CONNECT/ Capgemini eGovernment Benchmark

Indicator	A. Use of the Portal	Method
E.1	Overview of portal traffic statistics: number of unique visitors, geographic coverage	PSI Questionnaire
E.2	Top ten most accessed data sets.	PSI Questionnaire
E.3	Top ten stakeholders accessing the portal	PSI Questionnaire

Thus, to meet the goal of the landscaping and enable tailored support to be brought to the countries, the PSI expert group is kindly invited to fill in a questionnaire. 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. The picture below shows where the input from the landscaping questionnaire will be used in the maturity assessment.

		Sub-Indicators	Landscaping Questions
Open Data Maturity Assessment	Open Data Readiness	1. Presence of Open Data Policy	Section 2.1: Presence of Open Data policy Section 3: The use of Open Data Section 4: The impact of Open Data
		2. Licensing norms	Section 2.2: Licensing Norms
		3. Extent of central coordination	Section 2.3 : Extent of central coordination
		4. Transposition of the revised PSI directive	<i>Currently collected from ePSI</i>
	Data availability	5. Usability of the portal	<i>Mostly Automated</i>
		6. Use of Data	Section 3: The use of Open Data Section 4: The impact of Open Data
		7. Re-usability of data	<i>Automated</i>
		8. Spread of data across domains	Automated & OpenDataMonitor

Figure 4 Mapping of indicators in the PSI Questionnaire

2.2 Expected Output

The expected output of the landscaping will consist in an insight report presenting where each and every country stands on their Open Data Journey. Building on the monitoring activities conducted by Open Data Institute and the Capgemini benchmarking, European countries can roughly be clustered into three different levels of maturity as part of their efforts on a full-fledged open data journey. The level of maturity will then 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 data sets.

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, country fiches depicting the specific situation of each country will be drafted. Country fiches will also contain further information with regard to the impact, best practices and main barriers faced for further publication of data and its re-use. (See sections 4, 5 and 6 from the questionnaire.) The country fiches will therefore assess the political context in support to Open Data and the level of maturity and availability of data at national level.

The country fiches will follow a practical hands-on format as outlined in the draft below:

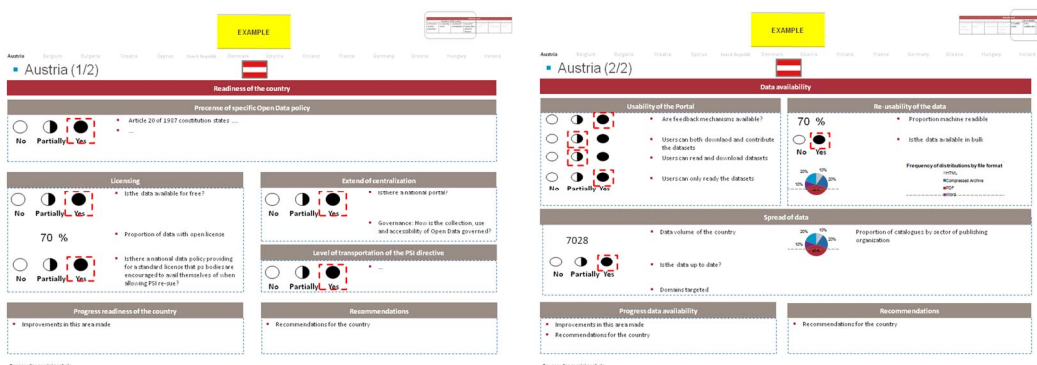


Figure 5 Draft Format Country Fiches

The country fiches will be updated on a yearly basis, thus contributing to at least three editions.

2.3 Timing of Work

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



Figure 6 Timing of Work

Annex I : Domains listed in the G8 Open Data Charter

To adopt an internationally recognised approach in structuring data sets, the consortium uses as reference the domains listed in the G8 Open Data Charter⁹.

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

⁹ https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/207772/Open_Data_Charter.pdf