European Data Portal

Measuring open data maturity

Sixth edition, 2020

The purpose of this document is to present an overview of the yearly open data maturity assessment adopted by the European Data Portal (EDP).

Table of Contents

1 Background and objective of the landscaping exercise

Since its launch in 2015, the European Data Portal (EDP) has been the main point of access at EU level to find public sector information published across Europe. The EDP covers a broad range of activities such as deploying, maintaining, and updating the EDP infrastructure as well as providing learning material and research reports on various open data topics. Moreover, it assists the national open data teams in the form of maturity assessments, tailored support, and strategic advice activities to enable their successful open data transformation.

The EDP objective is to improve access to open data, foster high-quality open data publication at national, regional, and local level, and increase its impact. Within this remit, the European Data Portal has been conducting an annual landscaping exercise providing the EU27 countries as well as the EFTA countries Iceland, Norway, Liechtenstein, and Switzerland with an assessment of their maturity level and documenting their year-on-year progress since 2015. The United Kingdom is also part of the assessment and this year, the assessment also includes several of the Eastern European Partnership countries, Moldova, Azerbaijan, and Ukraine. The landscaping exercise offers a benchmarking and learning tool at both national as well as European level. It supports countries to better understand their level of maturity, to capture their progress and the areas for improvement, and benchmark this against other countries. At the same time, the landscaping provides the evidence on which both generic and targeted support activities for the EU Member States is based.

In the period 2015-2017 the annual open data maturity measurement was built on two key indicators: “readiness” and “maturity”, covering the policy developments at country level as well as the level of sophistication of the national open data portals. To better reflect the open data developments taking place across Europe, a major update to the landscaping methodology was carried out in 2018. The 2018 methodology made the assessment more comprehensive and set a stronger focus on the quality of open data as well as the re-use and impact derived by open data. The scope of the assessment has hence been broadened to comprise four dimensions: policy, portal, impact, and quality.

In 2019, additional layers of granularity were added to the four dimensions. The updates to the questionnaire aims to provide further impulses for the national open data teams to redirect their focus on new strategic areas, such as stronger prioritisation of high-quality open data publication, an active fostering of re-use and monitoring mechanisms of open data re-use, the development of advanced portal features, and the need for more inclusive and participative governance structures.

The 2020 Open Data Maturity assessment has focused on the maintenance of the methodology in order to ensure continuity and enable comparison with previous years. Along this line of reasoning, the weighting of the dimensions and sub-dimensions remained the same as previous years, even though there may be reason to change it. This will be part of a more large-scale update to the questionnaire in 2021. Nevertheless, several minor adaptations to the survey questions have been implemented, to improve clarity or address ambiguities in response to the open data representatives’ feedback.

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.

Open data cannot be considered open if it is not accompanied by a licence that ensures its free re-use. Depending on the type of licences data is published under, the licence might stipulate that:

- Those who use the data must credit whoever is publishing it (this is called attribution)
- Those who combine the data with other data must 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 PSI Directives of 2003 and 2013 and the Open Data Directive of 2019[2] provide a common legal framework for a European market for government-held data (Public Sector Information).

3 Work approach

To perform the landscaping exercise, several activities are conducted on a yearly basis:

- *Step 1: Refine and update the landscaping methodology and questionnaire*
- *Step 2: Coordinate and assist national teams in filling out the landscaping questionnaire*
- *Step 3: Analyse and validate the data together with the national teams*
- *Step 4: Complement the results with additional desk research*
- *Step 5: Publish an in-depth report and country factsheets documenting the results and findings*
- *Step 6: Visualise the results on the dashboard of the European Data Portal*
- *Step 7: Produce an analytical report and webinars showcasing best practices from countries*

4 Open data dimensions

Similar to past iterations of this research, the data was collected through a questionnaire sent to the national open data representatives working in collaboration with the European Commission and the Public Sector Information Expert Group. The questionnaire was structured against the four open data dimensions as outlined below and included detailed metrics for each dimension to assess the level of maturity. Dimensions and metrics were last specified at the time of the latest major revision of the methodology in 2018. They have since been maintained to improve clarity or address ambiguities in response to the open data representatives' feedback.

Open Data Policy focuses on the presence of specific policies and strategies to foster open data at national level. The dimension also analyses the existence of governance structures that allow the

1 http://opendefinition.org/

2 Directive 2003/98/EC of 17 November 2003, Directive 2013/37/EU of 26 June 2013, Directive (EU)2019/1024 of 16 July 2019

participation of private and third sector actors, as well as implementation measures that enable open data initiatives at national, regional, and local level. Furthermore, the dimension looks at training schemes that enhance the data literacy skills of the civil servants working with data as well as harvesting mechanisms that foster the discoverability of all open data available in the country.

Open Data Portal focuses on advanced portal functions that enable both versed and less versed users to access open data via the national portal and features that enhance the interaction between publishers and re-users (via a forum or discussion boards). Additionally, the dimension assesses the extent to which portal managers use web analytics tools to better understand their users' needs and behaviour and update the portals' features in line with the insights gained from these analyses. The dimension examines the open data coverage across different domains, as well as the approach and measures in place to ensure the portal's sustainability.

Open Data Impact looks at the activities performed to monitor and measure open re-use and the impact derived by such re-use. Beyond this first layer of "strategic awareness", the impact dimension focusses on four areas of sectoral impact: political, social, environmental, and economic. Within these areas, the questionnaire examines the extent to which monitoring is in place to document the re-use of open data published in these fields, the extent to which applications, products, and services have been developed to address challenges in these fields, as well as the extent to which civil society initiatives exist that are based on such open data and supported by government institutions. With regard to the economic field, the questionnaire focuses on assessments, reports and/or studies that demonstrate the micro and macro-economic value of open data, as well as the efficiency gains achieved by the public sector.

Open Data Quality focuses on the measures adopted by portal managers to ensure the systematic harvesting of metadata from sources across the country, as well as the currency of the available metadata and where possible the actual data, the monitoring of the compliance with the DCAT-AP metadata standard as well as the quality of deployment of the published data. The fourth dimension provides impulses for portal managers and policy-makers to enable open data publication that is good quality all round: using open data formats, machine-readable, high-quality and suitable to a linked data approach (the use of URI's etc.).

5 Detailed dimensions and metrics

The indicators belonging to the four open data dimensions can be seen below.

Dimension 1: Open data policy	
Indicator	**Policy framework**
1.1	▪ Open Data policies and strategies are in place at national level. ▪ The open data policy/strategy has been updated in the past year. ▪ Opening PSI is mandatory at national level.
1.2	▪ The national open data strategy incentivises the re-use of open data by both the public and private sectors. ▪ The national open data policy/strategy incentivises access to real-time data. ▪ The development of data inventories in the public bodies at national, regional and local levels is defined as priority in the national policy and/or strategy.
1.3	▪ Priority domains for data release have been identified at national level. ▪ Data publication is prioritised in collaboration with other stakeholders (re-users).
Indicator	**Governance of open data**
2.1	▪ An open data governance structure is in place, that ensures the publication of open data at all government levels. ▪ The governance structure enables the development of open data initiatives at local and regional levels.
2.2	▪ Data stewards/ PSI officers are appointed at each public body level. ▪ A regular exchange between the data stewards as well as data publishers and re-users is ensured.
2.3	▪ Various open data events are organised throughout the country to foster the exchange on the open data topic. ▪ The events are organised by a mix of actors (public sector bodies at national, regional and local levels, private sector, civil society, universities).
Indicator	**Open data implementation**
3.1	▪ Guidelines are in place to assist publication at national, regional and local levels. ▪ Holders of dynamic data are assisted in providing real-time access to their data.
3.2	▪ Data publication plans exist at public body level and progress against these plans is monitored by the national level. ▪ Number of public bodies that still charge above marginal costs is monitored.
3.3	▪ Regional and local sources are systematically harvested by the national level. ▪ Percentage of harvested sources from total existing sources is known.
3.4	▪ Training activities for civil servants working with data are in place. ▪ The trainings provide a certification and/or are formally recognised as professional development for civil servants.

Dimension 2: Open data portal	
Indicator	**Portal feature**
1.1	▪ Portal features ensure discoverability and access to datasets and relevant content. ▪ Advanced features enable users to contribute content to the portal, provide feedback on existing content and rate featured datasets.

1.2	▪ The portal enables users to find information and news on relevant open data topics in the country.
1.3	▪ The portal enables the interaction and exchange between users and publishers as well as among re-users.
1.4	▪ Use cases are promoted via a designated section on the portal and mapped to the open data they are based on. ▪ Use cases can be submitted by different users to the portal.
1.5	▪ Preview functions to both tabular and geospatial data are available.
1.6	▪ Data publishers and other interested parties can find information on tools that enable the work with data.
Indicator	**Portal usage**
2.1	▪ The portal is responsive via both mobile and desktop.
2.2	▪ Traffic to the portal (number of unique visitors, visitor profiles, percentage of machine traffic etc.) is monitored by the portal managers.
2.3	▪ Analytics tools are used to derive insights into users' behaviour and needs. ▪ These insights are embedded into the portal update cycles.
2.4	▪ The most and least consulted categories and datasets are known. ▪ The most used search keywords are known, and updates are performed to ensure a better discoverability of available content.
2.5	▪ API usage is monitored and used to gain insights into user profiles (e.g. "power users").
Indicator	**Data provision**
3.1	▪ The majority of data publishers can contribute data to the national portal. ▪ Data publishers that do not contribute to the national portal have been identified and actions taken to enable data publication from these sources.
3.2	▪ Access to real-time data is provided via the portal. ▪ The percentage of real-time data from total featured data is known.
3.3	▪ A separate section exists on the portal where community-sourced data can be uploaded.
Indicator	**Portal sustainability**
4.1	▪ A sustainability strategy for the portal has been defined. ▪ Measures are in place to ensure the portal reaches its targeted audience.
4.2	▪ The portal team helps enhance the visibility of the portal and the featured datasets by organising/attending info sessions and/or events to promote the national portal.
4.2	▪ The source code and relevant documentation are available for the interested public. ▪ The national portal has accounts and an active presence on social media platforms.
4.3	▪ User satisfaction surveys on the national portal are regularly performed. ▪ The portal undergoes regular reviews. Users' feedback is considered in this review process.
4.4	▪ The portal provides a dashboard with information on the main performance indexes. ▪ The dashboard enables data publishers to analyse their own publication activities and quality.

Dimension 3: Open data impact	
Indicator	**Strategic awareness**
1.1	▪ Re-use of open data is monitored at national level via, for example, the national portal.
1.2	▪ Activities are in place at public body level to boost and monitor the re-use of own published data.
1.3	▪ Activities are in place at national level to measure the re-use of open data. ▪ A methodology to measure the impact of open data is in place or first steps in this direction are taken.
Indicator	**Political impact**
2.1	▪ Activities have been launched to monitor the political impact of open data.
2.2	▪ Various re-use examples exist that showcase the impact of open data on increasing government efficiency and effectiveness.
2.3	▪ Various re-use examples exist that showcase the impact of open data on increasing transparency and accountability.
2.4	▪ Various re-use examples exist that showcase the impact of open data on enable better policy and decision-making processes.
2.5	▪ Civil society initiatives that are open data driven and aim to tackle a challenge identified in the political field exist / are supported by government.
Indicator	**Social impact**
3.1	▪ Activities have been launched to monitor the social impact of open data.
3.2	▪ Various re-use examples exist that showcase the impact of open data on a better inclusion of marginalised groups.
3.3	▪ Various re-use examples exist that showcase the impact of open data on increasing awareness on housing issues.
3.4	▪ Civil society initiatives that are open data driven and aim to tackle a challenge identified in the social field exist / are supported by government. ▪ Studies have been conducted that deal with the social impact of open data.
Indicator	**Environmental impact**
4.1	▪ Activities have been launched to monitor the environmental impact of Open Data.
4.2	▪ Various re-use examples exist that showcase the impact of open data on increasing awareness on air and water quality.
4.3	▪ Various re-use examples exist that showcase the impact of open data on increasing awareness on noise levels in cities.
4.4	▪ Various re-use examples exist that showcase the impact of open data on enabling better waste management and increasing awareness on waste reduction.
4.5	▪ Civil society initiatives that are open data driven and aim to tackle a problem identified in the environmental exist / are supported by government. ▪ Studies have been conducted that deal with the environmental impact of open data.
Indicator	**Economic impact**
5.1	▪ Activities have been launched to monitor the economic impact of open data.
5.2	▪ Various re-use examples exist that showcase the impact of open data on a macro-economic level.

5.3	▪ Various re-use examples exist that showcase the impact of open data on a micro-economic level.
5.4	▪ Civil society initiatives that are open data driven and aim to tackle economic problems exist / are supported by government. ▪ Studies have been conducted that deal with the economic impact of open data.

Dimension 4: Open Data Quality	
Indicator	**Currency and completeness**
1.1	▪ A pre-defined approach is in place to ensure metadata is up to date.
1.2	▪ Harvesters are programmed to ensure that changes at the source are reflected with the least amount of delay on the national portal.
1.3	▪ The portal provides access to a vast range of data, both historical and current.
Indicator	**Monitoring and measures**
2.1	▪ Mechanisms are in place to monitor the quality of the metadata. ▪ Information on the quality of the metadata is available to the broader public.
2.2	▪ Guidelines and/or tools are available to assist publishers in choosing the right type of licence for their data. ▪ The compliance level in terms of correct licencing information is monitored.
2.3	▪ Measures are in place to assist publishers in publishing in high-quality metadata / data.
Indicator	**DCAT-AP Compliance**
3.1	▪ Guidelines and materials to help publishers in ensuring compliance with DCAT-AP are linked on the national portal.
3.2	▪ Compliance with the DCAT-AP standard in terms of mandatory, recommended and optional classes is monitored. The compliance level is monitored.
3.3	▪ Monitoring activities of the percentage of accessible distributions (availability of AccessURL and DownloadURL) are in place. The compliance level is monitored.
Indicator	**Deployment quality and linked data**
4.1	▪ A model (such as the 5-star open data model or similar) is used to assess the quality of data deployment. ▪ Activities to familiarise publishers with this model and linked data are conducted.
4.2	▪ The percentage of published open data that complies with the requirements of the 1, 2, 3, 4 and 5-stars deployment levels is known. ▪ The improvements in terms of quality of open data deployment are monitored.

6 Scoring

The scoring of the open data maturity assessment can be seen below.

Dimension	Key metrics	Scoring	Weight
Open data policy		**650**	**25%**
	Policy framework	220	
	Governance of open data	220	
	Open data implementation	210	
Open data Portal		**650**	**25%**
	Portal features	240	
	Portal usage	150	
	Data provision	110	
	Portal sustainability	150	
Open data impact		**650**	**25%**
	Strategic awareness	140	
	Political impact	130	
	Social impact	120	
	Environmental impact	150	
	Economic impact	110	
Open Data Quality		**650**	**25%**
	Currency and completeness	150	
	Monitoring and measures	150	
	DCAT-AP compliance	180	
	Deployment quality and linked data	170	
Total		**2600**	**100%**

7 Output

The output of the annual open data maturity assessment[3] consists of:

- **The open data maturity report** supports countries to better understand their level of maturity, to capture their progress and the areas for improvement, and benchmark this against other countries. Additionally, the report provides an overview of best practices implemented across Europe that could be transferred to other national and local contexts.
- **The open data maturity dashboard** presents the detailed country scores per dimension as well as visualisations of the maturity levels across Europe.
- **The customised country factsheets** provide a more detailed insight at national level into the results of the four open data dimensions in comparison with EU27 average and the results from previous years.

[3] All output material from the 2020 open data maturity assessment and previous editions can be found at: https://www.europeandataportal.eu/en/impact-studies/open-data-maturity