

Measuring Open Data Maturity

Ninth edition, 2023

This document describes the methodology of the annual open data maturity assessment conducted by data.europa.eu.

Licence: CC-BY



Contents

1.	Objectives of the open data maturity assessment	1
2.	History of the open data maturity assessment	1
3.	Working definitions and legislative background	2
4.	Work approach	3
5.	Open data dimensions	4
6.	Indicators and metrics	6
7.	Scoring	12
8.	Outputs	13

Objectives of the open data maturity assessment

Since its launch in 2015, data.europa.eu (¹) has been the main point of access at the EU level to public sector information published across Europe. data.europa.eu aims to improve access to open data, as well as to foster both high-quality open data publication and the reuse of open data to create impact.

Within this remit, data.europa.eu conducts an annual landscaping exercise for European countries on their open data maturity (ODM). Participation is voluntary, and the scope of the assessment includes the Member States of the European Union (the EU-27) (²), the member states of the European Free Trade Association (³) and the candidate countries for EU membership (⁴). The 2023 ODM assessment includes 35 participating countries: the EU-27, 3 European Free Trade Association countries (Iceland, Norway and Switzerland), and 5 candidate countries (Bosnia and Herzegovina, Montenegro, Albania, Serbia and Ukraine).

The purpose of the ODM assessment is to evaluate the development of countries in making public sector information available and stimulating its reuse. The landscaping exercise offers a benchmarking and learning tool for use at both the national and European levels. The results of the assessment support countries in better understanding their relative level of maturity compared with other countries and their year-on-year progress and in identifying further areas for improvement. The landscaping exercise provides evidence and recommendations on which activities European countries could adopt to further their ODM.

2. History of the open data maturity assessment

The first three editions of the ODM (2015–2017) used two dimensions to assess ODM: (1) open data readiness and (2) portal maturity, which evaluated policy developments at the national level and the degree of sophistication of national open data portals, respectively. In 2018, a major update to the landscaping methodology was carried out to better reflect open data developments across Europe. This revision of the methodology made the assessment more comprehensive and placed a stronger emphasis on the quality of open data and the reuse of and impact derived from open data. The scope of the evaluation was broadened to cover four dimensions: (1) policy, (2) impact, (3) portal and (4) quality.

In 2019, additional layers of granularity were added to the four dimensions. The updates to the assessment aimed to provide further stimulus for the national open data teams to redirect their focus onto new strategic areas – such as greater prioritisation of high-quality open data publication, active fostering of mechanisms to monitor open data reuse, and the development of advanced portal features such as multifaceted search and user feedback functionalities – and to raise awareness of the need for more inclusive and participative governance structures.

In 2022, the methodology underwent another structured revision to further challenge European countries to increase their ODM, as a means of systematically stimulating ODM to grow beyond the

^{(1) &}lt;u>data.europa.eu</u> is the official portal for European open data. The portal was launched in 2021, formed from the merger of the European Data Portal and the European Union Open Data Portal into a single coherent core component of the public sector data infrastructure that has been set up by the European Union, its institutions and the Member States.

^{(2) &}lt;a href="https://european-union.europa.eu/principles-countries-history/country-profiles-en">https://european-union.europa.eu/principles-countries-history/country-profiles-en.

^{(3) &}lt;a href="https://www.efta.int/">https://www.efta.int/.

^{(4) &}lt;a href="https://european-union.europa.eu/principles-countries-history/joining-eu en.">https://european-union.europa.eu/principles-countries-history/joining-eu en.

current level and of pushing countries to keep pace with policy changes in the field, such as the proposals on high-value datasets. To this end, all four dimensions and related questions were reviewed. Across the four dimensions, questions were streamlined to better include initiatives at the regional and local levels and specific types of open data. In addition, a focus was introduced on countries' level of preparedness for the European Commission's upcoming implementing regulation on high-value datasets (5). A major change in the 2022 methodological update was the restructuring of the impact dimension. This was done to better acknowledge the challenge that countries face in assessing the impact of open data and to better distinguish between measuring reuse of open data and measuring the impact created through that reuse. This involved adding a new indicator, on measuring impact, to the impact dimension.

The 2023 edition of the methodology does not include major methodological updates. The wording of the questions underwent a general review for clarity, especially those related to high-value datasets, given the publication of Implementing Regulation (EU) 2023/138 (6) in January 2023. The implementing regulation will be applicable from June 2024, and the questions related to high-value datasets aim to track the preparations made by EU Member States to implement the regulation's requirements. European Free Trade Association countries and candidate countries can choose 'not applicable' when answering questions regarding specific EU legislative provisions and still be awarded full points in the scoring schema.

The next structural revision of the methodology is planned for 2024.

3. Working definitions and legislative background

This section provides a working definition of what should be understood by 'open data'.

'Open data', when referring to data issued by government, means public sector data and publicly funded data made available for reuse for any purpose (7).

Open data must be accompanied by a licence that ensures its free reuse. The licence may stipulate that:

- those who use the data must credit the publisher ('attribution');
- those who transform the data, including by combining it with data from other sources, must release the output under the same (open) licence ('share alike').

These and other principles on open data are described in detail on the Open Definition website (8).

The ODM assessment is informed by the EU's open data policies, primarily Directive (EU) 2019/1024 on open data and the reuse of public sector information (which entered into force in July 2019 and was to be transposed into national law by July 2021) (9), which serves as a stimulus to the reuse of open data and creates the legal basis for the activities carried out by national open data teams. The open data directive 'provides common rules for a European market for government-held data'. This

^{(5) &}lt;a href="https://digital-strategy.ec.europa.eu/en/news/commission-defines-high-value-datasets-be-made-available-re-use">https://digital-strategy.ec.europa.eu/en/news/commission-defines-high-value-datasets-be-made-available-re-use.

^{(6) &}lt;a href="https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=uriserv%3AOJ.L">https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=uriserv%3AOJ.L .2023.019.01.0043.01.ENG.

^{(7) &}lt;a href="https://data.europa.eu/en/dataeuropa-academy/what-open-data">https://data.europa.eu/en/dataeuropa-academy/what-open-data.

^{(8) &}lt;a href="http://opendefinition.org/">http://opendefinition.org/.

⁽⁹⁾ https://eur-lex.europa.eu/legal-content/EN/TXT/?qid=1561563110433&uri=CELEX:32019L1024.

directive replaces the public sector information directive (¹⁰), which aimed to make public sector data (i.e. from public sector bodies in EU Member States, at the national, regional and local levels) and publicly funded data (e.g. from meteorological institutes) reusable.

The open data directive also mandated the adoption of an implementing act on high-value datasets, and the regulation discussed above was published in January 2023 (11). High-value datasets are select public sector datasets identified as having important benefits for society, the environment and the economy. They fall into six categories: geospatial, earth observation and environment, meteorological, statistics, companies and mobility. High-value datasets must be made available for reuse with minimal legal and technical restrictions and free of charge. In addition, these datasets must be provided through application programming interfaces (APIs) and, where relevant, as a bulk download.

The ODM assessment also includes questions about data that cannot be made open (e.g. data covered by the Data Governance Act (¹²), which came into force in September 2023) because it is necessary to have an overview of such data when making publication plans.

4. Work approach

The data for ODM is collected through a voluntary questionnaire sent to national open data representatives. This is done in collaboration with the European Commission and the Public Sector Information Expert Group. The questionnaire consists of self-assessment questions that the representatives of the national open data teams answer. Most questions have a predefined list of answer options (e.g. 'Yes' or 'No') from which the respondent selects an appropriate response. In addition, most questions request additional supporting information on the answer chosen, such as a URL link to relevant material or a description of related activities. Following completion, the responses are sent to the research team at data.europa.eu.

Once the research team receives the questionnaires, they undergo three rounds of validation. The first round involves a high-level check in which the research team checks each questionnaire for completeness. Following this, countries are given the opportunity to provide input on any missing answers. The second and third rounds of validation involve an in-depth review of the questionnaires, in which two reviewers from the research team (one for round two and one for round three) independently check the responses. The reviewers assess whether the explanations accompanying the answers are complete, relate to the question and sufficiently justify the answer option selected. Throughout the second and third validation rounds, the reviewers may refer to previous ODM questionnaires from the countries or perform desk research for additional clarification. The reviewers mark insufficiently answered questions that require further input from the countries. After the second validation round, the comments from the first reviewer are shared with the countries, who are invited to modify their responses to address the reviewer's comments. After the third validation round, the comments from the second reviewer are shared alongside the first reviewer's comments, and the countries are invited to address the remaining flagged questions. The research team finalises the scores based on the responses to the flagged questions.

^{(10) &}lt;a href="https://eur-lex.europa.eu/legal-content/en/ALL/?uri=CELEX%3A32003L0098">https://eur-lex.europa.eu/legal-content/en/ALL/?uri=CELEX%3A32003L0098.

^{(11) &}lt;a href="https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=uriserv:OJ.L..2023.019.01.0043.01.ENG">https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=uriserv:OJ.L...2023.019.01.0043.01.ENG.

^{(12) &}lt;a href="https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32022R0868">https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32022R0868.

5. Open data dimensions

As stated previously, the ODM questionnaire is structured in four dimensions: (1) open data policy, (2) open data impact, (3) open data portal and (4) open data quality. These dimensions are subdivided into indicators. The four dimensions are outlined below.

Open data policy investigates countries' policies and strategies regarding open data, the national governance models for managing open data and the measures deployed to implement the policies and strategies. To evaluate these elements, the dimension is composed of three indicators: (1) policy framework, (2) governance of open data and (3) open data implementation.

Indicator	Key elements
Policy framework	An open data policy and strategy are in place at the national level to provide a long-term strategic vision and action plan for open data. The strategies incentivise open data reuse in both the public and private sectors, as well as access to real-time, geospatial and citizen-generated data. Preparatory activities regarding high-value datasets are in place. Open data policies and strategies align with the European Commission's priorities for 2019–2024 (¹³).
Governance of open data	Governance models and regular coordination activities across public sector bodies are in place to ensure the publication of open data at all government levels and to support local and regional open data initiatives. Regular exchanges occur between open data providers and reusers from academia, businesses and non-governmental organisations.
Open data implementation	Data publication plans and implementing processes exist. The number of public bodies that charge above the marginal costs of dissemination for the reuse of their open data is monitored. Training activities for civil servants working with data are organised, along with society-wide open data literacy initiatives.

Open data impact assesses the whether countries have defined reuse and the extent to which they are prepared to measure it, the actions taken by countries to measure reuse and understand the needs of reusers, and the presence of examples of reuse cases in the domains of government, society, the environment and the economy. To evaluate these elements, the dimension is composed of three indicators: (1) strategic awareness, (2) measuring reuse and (3) created impact.

Indicator	Key elements
	There is a national definition of open data reuse. Mechanisms are in place at the national, regional or local level to monitor and foster open data reuse, including in relation to high-value datasets. A methodology exists to measure the impact derived from reusing open data.
Measuring reuse	Tools are in place to understand which datasets are reused and how. There is a process for gathering and classifying reuse cases systematically. Activities are performed to better understand reusers' needs.

⁽¹³⁾ https://commission.europa.eu/strategy-and-policy/priorities-2019-2024 en.

Created impact

- governmental
- social
- environmental
- economic

The impact created by open data has been systematically studied, and reuse examples exist that showcase the impact of open data in the governmental, social, environmental and economic domains.

Open data portal investigates the functionality of national open data portals, the extent to which users' needs and behaviour are examined to improve the portal, the availability of open data across different domains and the approach to ensuring the portal's sustainability. To evaluate these elements, the dimension is composed of four indicators: (1) portal features, (2) portal usage, (3) data provision and (4) portal sustainability.

Indicator	Key elements
Portal features	Portal features ensure access to datasets and relevant content, and include more advanced features such as SPARQL search, discussion forums, rating of datasets, requesting datasets and providing transparency on the progress status of requested datasets. Activities to promote the visibility and reuse of high-value datasets through the portal are planned.
Portal usage	Traffic to the portal is monitored, and analytics tools are used to gain insights into users' behaviour and the most and least consulted data categories. In addition, the portal offers APIs through which advanced users can access the metadata programmatically.
Data provision	Most data providers contribute data to the national portal, and actions are taken to enable publication from data providers. In addition, access to real-time data is provided through the portal, and data that does not stem from official sources can be uploaded. Furthermore, data from regional or local sources is discoverable on the national portal.
Portal sustainability	A strategy to ensure the sustainability of the portal has been determined, and activities are conducted to ensure the portal's visibility, including through a social media presence. In addition, user surveys are conducted regularly and feed into a review process to improve the portal.

Open data quality assesses the measures adopted by portal managers to ensure the systematic and timely harvesting of metadata and the monitoring mechanisms in place to ensure the publication of metadata compliant with the DCAT-AP metadata standard and several deployment quality requirements. To evaluate these elements, the dimension is composed of four indicators: (1) metadata currency and completeness, (2) monitoring and measures, (3) DCAT-AP compliance and (4) deployment and linked data.

Indicator	Key elements
Metadata currency and completeness	A systematic approach is in place to ensure that metadata is up to date. Programmes that harvest metadata automatically are used to ensure that changes at the source are reflected with a minimum of delay on the national portal. The portal provides access to a vast range of data, both historical and contemporary. Preparations are under way to ensure that high-value data is interoperable with high-value datasets from other countries.
Monitoring and measures	Mechanisms are in place to monitor metadata quality on the national portal, as well as its compliance with licensing standards. Measures are in place to assist data providers in publishing high-quality metadata and choosing the right type of licence for their data.
DCAT-AP compliance	Compliance with the DCAT-AP standard regarding mandatory, recommended and optional classes is monitored. Guidelines and learning materials help data providers in ensuring compliance with DCAT-AP.
Deployment quality and linked data	A model is used to assess the quality of data and metadata deployment. The percentage of published open data that complies with specific deployment quality requirements including having links to other data sources is known, and improvements in terms of deployment are monitored.

6. Indicators and metrics

The indicators within each dimension are assessed through several questions that pertain to specific concepts. The tables below provide a summary of the key concepts assessed for each indicator.

	Dimension 1: Open data policy	
1.1. Policy fr	1.1. Policy framework	
1.1.1.	 Open data policies and strategies are in place at the national, regional and/or local levels. 	
	The open data policies/strategies have been updated in the past 2 years.	
	The open data directive is implemented at the national level.	
1.1.2.	 The (national) open data strategy incentivises the public and private sectors to reuse open data. 	
	 The (national) open data policies/strategies incentivise access to real-time and dynamic data, citizen-generated data and geospatial data. 	
	 The (national) open data policies/strategies incentivise the development of data inventories in national, regional and local public bodies. 	
1.1.3.	 Preparatory activities to identify high-value data domains to be prioritised for publication are in place. 	
	 Measures are in place to facilitate other stakeholders' involvement in this prioritisation process. 	

	 Progress has been made to ensure that public bodies holding high-value datasets are prepared to denote those datasets as such in their metadata. 		
1.1.4.	 Open data policies/strategies align with the EU's overarching priorities. 		
1.2. Governa	1.2. Governance of open data		
1.2.1.	 An open data governance structure that ensures open data publication at all government levels is in place. 		
	 The governance structure enables the development of open data initiatives at the local and regional levels. 		
1.2.2.	 Details of the person or team responsible for open data activities in the country are publicly available. 		
	 A document describing the responsibilities and working approach of the national open data team (and possibly those of regional and/or local teams) is publicly available. 		
	 Regular exchanges between the national open data team and the team maintaining the national and/or local portal(s) are ensured. 		
1.2.3.	 Open data officers have been appointed at each public body level. 		
	 Regular exchanges between the national open data team and open data officers are ensured. 		
	 Regular exchanges between open data officers, data providers and reusers are ensured. 		
1.3. Open d	ata implementation		
1.3.1.	 Data publication plans exist at the public body level, and progress against these plans is monitored at the national level. 		
	 The number of public bodies still charging above marginal costs is monitored. 		
1.3.2.	 Measures are in place to address the challenges faced in implementing the aforementioned open data policies/strategies. 		
	 Holders of dynamic data are assisted in providing real-time access to their data. 		
	 Holders of geospatial data are assisted in making this data publicly available. 		
	 Citizens/organisations are assisted in providing access to citizen-generated data. 		
1.3.3.	 Training activities for civil servants working with (open) data are in place. 		
	 Training activities result in certification and/or are formally recognised as professional development for civil servants. 		
	 Society-wide open data literacy initiatives are in place. 		

	Dimension 2: Open data impact
2.1. Strategi	c awareness
2.1.1.	 Reuse of open data is monitored at the national, regional or local level, for example through the national portal. This includes monitoring reuse of high- value datasets.
2.1.2.	 Activities are in place at the public body level to boost and monitor the reuse of bodies' own published data.
2.1.3.	A definition of reuse is in place.
	A methodology to measure the impact of open data is in place.
2.2. Measurii	ng reuse
2.2.1.	 Activities are in place to understand which datasets are reused and how, for example: log files are analysed to collect information on and investigate how datasets are being used; automated feedback mechanisms are in place to track users' access to datasets; Surveys / other extensive research tools are used to measure the reuse of open data.
2.2.2.	 Activities are in place to better understand reusers' needs, for example: feedback sessions with portal users are conducted regularly; social media sentiment analysis is used.
2.2.3.	 A process is in place to systematically gather reuse cases. Reuse cases are classified according to categories (e.g. environmental, social, economic, etc.)
2.3. Created	impact
2.3.1.	 Data on the impact created by open data on governmental challenges is available in the country.
	 Various reuse examples exist that showcase the impact of open data on:
	 increasing government efficiency and effectiveness in delivering public services; increasing the transparency and accountability of public administrations; enabling better policy- and decision-making.
2.3.2.	 Data on the impact created by open data on societal challenges is available in the country.
	Various reuse examples exist that showcase the impact of open data on:
	 better including marginalised groups and reducing inequality; raising awareness of urban housing issues; raising awareness of health- and well-being-related issues; raising awareness of education issues.

2.3.3.	 Data on the impact created by open data on environmental challenges is available in the country.
	Various reuse examples exist that showcase the impact of open data on:
	 raising awareness of biodiversity-related topics (e.g. air and water quality); enabling more environmentally friendly cities; raising awareness of climate change and connected disasters; encouraging lower energy consumption based on reducing fuel use and the switch to renewables.
2.3.4.	Data on the impact created by open data on the economy is available in the country.
	Various reuse examples exist that showcase the impact of open data on the following indicators of economic growth:
	 employment, technology and innovation, entrepreneurship and business creation.

Dimension 3: Open data portal			
3.1. Portal f	eatures		
3.1.1.	 Portal features ensure discoverability of and access to datasets (including through APIs) and relevant content. 		
	 Portal users can find documentation about using APIs and other tools that enable working with metadata, such as through search functionalities. 		
3.1.2.	 Advanced features enable users to provide content for the portal, give feedback on existing content and rate featured datasets. 		
3.1.3.	 The portal enables users to find information and news on relevant open data topics in the country. 		
3.1.4.	 The portal enables interaction and exchange between reusers and publishers. 		
3.1.5.	 Use cases are promoted through a designated section on the portal and mapped to the open data they are based on. 		
	 Reusers can submit use cases to the portal. 		
3.1.6	 Preview functions for both tabular and geospatial data are available. 		
3.1.7.	 Progress is being made on promoting the visibility and reuse of high-value datasets via the portal. 		
3.2. Portal u	3.2. Portal usage		
3.2.1.	The portal is responsive in both mobile and desktop versions.		
3.2.2.	 Traffic to the portal (e.g. number of unique visitors, visitor profiles, percentage of outgoing portal traffic generated through APIs', number of downloads as a 		

	proportion of number of datasets) is monitored by the portal team.
3.2.3.	 Analytics tools are used to derive insights into users' behaviour and needs.
	 These insights are embedded in the portal update cycles.
3.2.4.	The most and least consulted categories and datasets are known.
	 The most used search keywords are known, and updates are performed to ensure greater discoverability of available content.
3.2.5.	API usage is monitored and the results used to gain insights into user profiles.
3.3. Data p	rovision
3.3.1.	 Most data providers can submit data to the national portal.
	 Data providers that do not contribute to the national portal have been identified, and actions have been taken to enable data publication from these sources.
3.3.2.	Local or regional data sources are discoverable through the national portal.
	 Metadata from local or regional data sources is harvested automatically.
3.3.2.	 Access to real-time data is provided through the national portal.
	The percentage of real-time data in the total data featured on the portal is known.
3.3.3.	 A separate section exists on the portal where community-sourced/citizen- generated data can be uploaded.
3.4. Portal	sustainability
3.4.1.	 Measures are in place to ensure that the portal reaches its target audience.
	The national portal has accounts and an active presence on social media platforms.
	The portal team helps to enhance the visibility of the portal and the featured datasets by organising/attending information sessions and/or events to promote the national portal.

	Dimension 4: Open data quality	
4.1. Metadata currency and completeness		
4.1.1.	A predefined approach is in place to ensure that metadata is up to date.	
4.1.2.	 Mechanisms are in place to ensure that changes at the source are reflected with a minimum of delay on the national portal. 	
4.1.3.	The portal provides access to a vast range of historical and current data.	
4.1.4.	 Preparation is under way to ensure the interoperability of high-value datasets with those of other countries. 	
4.2. Monitori	ng and measures	
4.2.1.	 Mechanisms are in place to monitor the quality of metadata. 	
	 Information on metadata quality is available to the broader public. 	
4.2.2.	 Guidelines and/or tools are available to assist data providers in choosing the right licence for their data. 	
	■ The compliance level in terms of correct licensing information is monitored.	
4.2.3.	 Support (e.g. documentation, tools, a helpline) is in place to assist data providers in improving data quality. 	
4.3. DCAT-AP	compliance	
4.3.1.	 Guidelines and materials to help data providers ensure compliance with DCAT- AP are available on the national portal. 	
4.3.2.	 Compliance with the DCAT-AP standard regarding mandatory, recommended and optional classes is monitored. 	
4.3.3.	 Monitoring activities on the percentage of accessible distributions (i.e. the availability of 'access URL' and 'download URL' properties) are in place. 	
4.4. Deploym	ent quality and linked data	
4.4.1.	 A model (e.g. the 5-star open data model or similar) is used to assess the quality of data deployment. 	
	 Activities are in place to familiarise data providers with this model and linked data. 	
4.4.2.	 The percentage of published open data that complies with the chosen quality model is known. 	
	 Improvements in terms of the quality of open data deployment are monitored. 	

7. Scoring

Countries are scored on a list of questions relating to each indicator. Each question-and-answer selection is worth a different number of points (see the questionnaire for a full breakdown). Where relevant, choosing 'not applicable' as an answer is worth full points, such as when EU legislation does not apply to a country. The scores for the individual questions sum together to provide a total score for the indicator. In turn, the indicator scores are added together to give scores for the dimensions. The maximum scores for the indicators and dimensions are shown in the table below. The overall maturity score is calculated as the weighted percentage of all the dimensions, meaning that each dimension contributes an equal 25 % towards the overall maturity score.

Dimension	Indicators	Maximum score
Policy		640
	Policy framework	275
	Governance of open data	190
	Open data implementation	175
Impact		600
	Strategic awareness	170
	Measuring reuse	110
	Created impact:	320
	Governmental	80
	Social	80
	Environmental	80
	Economic	80
Portal		650
	Portal features	265
	Portal usage	145
	Data provision	100
	Portal sustainability	140
Quality		650
	Metadata currency and completeness	130
	Monitoring and measures	160
	DCAT-AP compliance	180
	Deployment quality and linked data	180
Total maximum score		2 540

8. Outputs

The outputs of the annual ODM assessment (14) consist of:

- a report that discusses the results of the assessment and trends in the collected data, while also
 providing an overview of best practices implemented across Europe that could be transferred
 to other national and local contexts;
- o country factsheets, which provide a maturity overview of each participating country;
- o a results overview that summarises the processed data;
- o the completed questionnaires, which hold the raw data.

^{(14) &}lt;a href="https://data.europa.eu/en/impact-studies/open-data-maturity">https://data.europa.eu/en/impact-studies/open-data-maturity.