

Quality

2024
Open Data Maturity Report

This study has been prepared as part of data.europa.eu. data.europa.eu is an initiative of the European Commission and is the official portal for European open data. The Publications Office of the European Union manages data.europa.eu.

European Commission

Directorate-General for Communications Networks, Content and Technology
Unit G.1 – Data Policy and Innovation
Email: CNECT-G1@ec.europa.eu

data.europa.eu

Email: info@data.europa.eu

Authors – Capgemini Invent

Martin Page PhD
Arman Behrooz
Maddalena Moro

Last updated: 10 March 2025

Disclaimer

The views set out in this publication are those of the authors and do not necessarily reflect the official opinion of the Commission. The Commission does not guarantee the accuracy of the data included in this study. Neither the Commission nor any person acting on the Commission's behalf may be held responsible for the use that may be made of the information contained therein.

Luxembourg: Publications Office of the European Union, 2024

© European Union, 2024



The Commission's reuse policy is implemented under Commission Decision 2011/833/EU of 12 December 2011 on the reuse of Commission documents (OJ L 330, 14.12.2011, p. 39, ELI: <http://data.europa.eu/eli/dec/2011/833/oj>). Unless otherwise noted, the reuse of this document is authorised under the Creative Commons Attribution 4.0 International (CC BY 4.0) licence (<https://creativecommons.org/licenses/by/4.0/>). This means that reuse is allowed, provided appropriate credit is given and any changes are indicated.

Chapter 6: Open data quality

The quality of the data refers to the overall state of the dataset. Preparing high-quality data includes dealing with missing values and other inaccurate elements, harmonising data structures and making the data available in accessible formats. Data quality also depends on the quality of its deployment on national portals, which can be assessed by looking at the use of aspects such as open data licences, machine-readable data formats, unique resource identifiers (a character sequence that identifies a dataset) and a linked data approach (a set of design principles for relating datasets to one another).

In addition to the data itself, high-quality data is accompanied by good descriptions. Such descriptive data is called metadata and gives information about other data, such as author, date and keywords. Specifications such as the Data Catalog Vocabulary – Application Profile ([DCAT-AP](#)) – which was designed to describe public sector datasets in Europe and is, therefore, the reference specification in the open data maturity (ODM) assessment methodology – define the structure and content of metadata descriptions and aim to make public sector data more easily searchable across borders and sectors.

Data that is high quality has greater value. This value derives from characteristics such as being easier for reusers to analyse and visualise. High-quality metadata similarly aids reuse by making datasets more discoverable, since search engines can better match the data's description with a user's search terms.

The **quality** dimension of the ODM assessment encourages national portals to publish datasets with high-quality data and metadata. The ODM methodology emphasises metadata quality, since national portals aim to make datasets discoverable and harvest metadata. The methodology also investigates whether portal managers have materials and processes to assist and incentivise data publishers to provide high-quality data.

In brief, the **quality** dimension 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 that is compliant with the DCAT-AP metadata standard and several deployment quality requirements. Table 1 summarises the key elements of the quality dimension.

Table 1: Indicators of the quality dimension

| 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 minimal delay on the national portal. The portal provides access to a vast range of historical and contemporary data. Preparations are under way to ensure that high-value data is interoperable with high-value datasets (HVDs) from other countries. |
| Monitoring and measures | Mechanisms are in place to monitor metadata quality on the national portal and 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. |

This chapter will first present overall performance on the policy dimension and then provide a summary of the results and best practices for each indicator.

Contents

| | |
|---|-----------|
| 6.1. Overall performance on the quality dimension | 5 |
| 6.2. Metadata currency and completeness | 8 |
| Currency of metadata | 8 |
| Completeness of metadata..... | 11 |
| Interoperability of high-value datasets | 12 |
| 6.3. Monitoring and measures..... | 13 |
| Monitoring the quality of metadata on portals..... | 13 |
| Setting metadata standards and licensing requirements..... | 16 |
| Support activities for data providers | 17 |
| 6.4. DCAT-AP compliance | 20 |
| Creating a framework for DCAT-AP compliance..... | 20 |
| Compliance with the DCAT-AP specifications..... | 21 |
| 6.5. Deployment quality and linked data | 23 |
| Use of models for deployment quality | 23 |
| Activities for data providers to ensure high-quality data | 24 |
| Deployment quality | 25 |
| 6.6. Pilot indicator: automated tests of metadata quality..... | 26 |

6.1. Overall performance on the quality dimension

The quality dimension is the least mature dimension of the ODM assessment according to the EU-27 average in 2024 (Figure 1). The average maturity of EU Member States in the quality dimension is 79.7%. This is a 3 percentage point (pp) decrease from 2023, primarily driven by a 6 pp decrease in the ‘monitoring and measures’ indicator and a 3 pp decrease in the ‘DCAT-AP compliance’ indicator. These decreases may be attributed to the introduction of 11 new questions or criteria related to the quality dimension in this year’s questionnaire, which introduced a higher set of requirements that were not previously measured. In addition, several countries reported lower metadata quality scores on the same questions asked last year.

Quality maturity score over time

EU-27, 2018–2024

- Quality dimension
- △ Metadata currency and completeness
- ◇ Monitoring and measures
- ◇ DCAT-AP compliance
- × Deployment quality and linked data

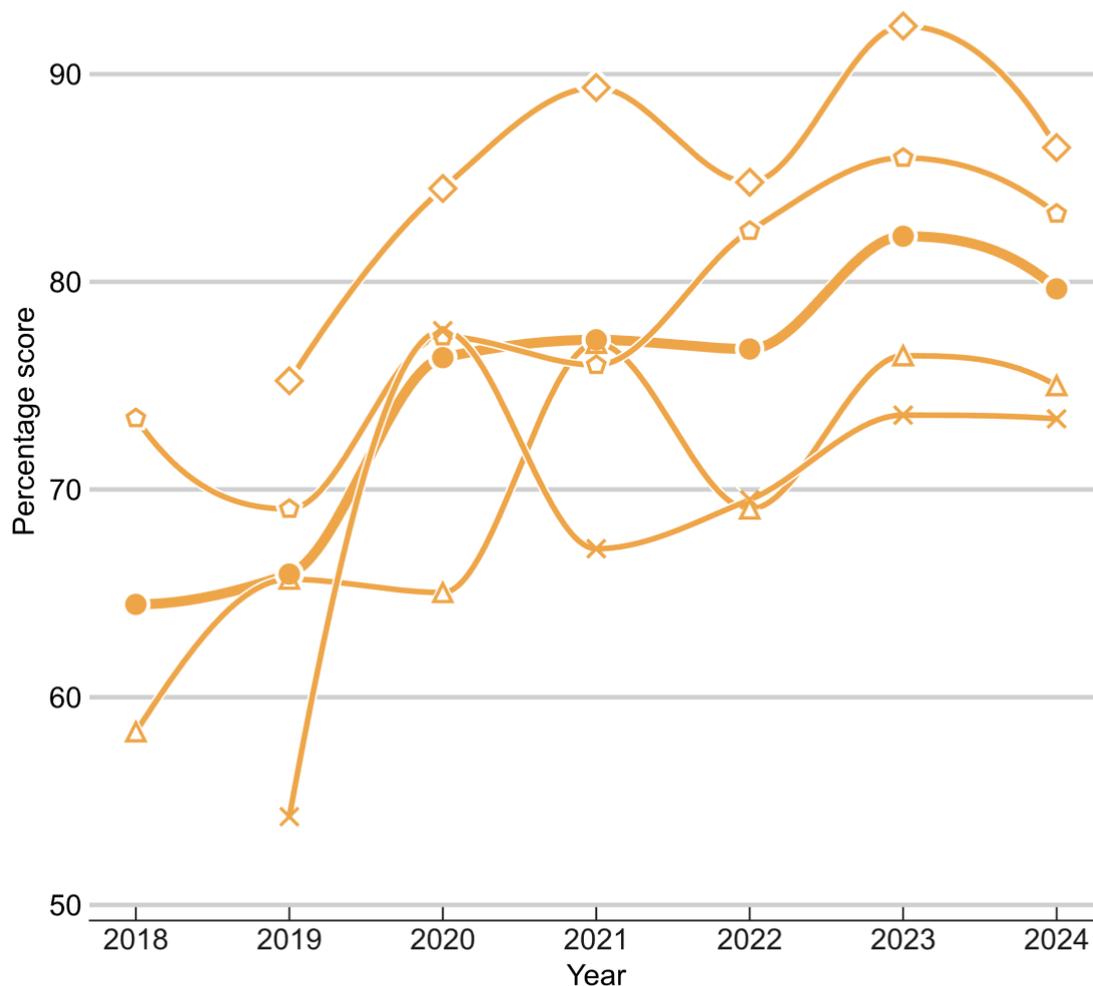


Figure 1: The EU-27 average score on the quality dimension decreased year-on-year but is still more mature than in years before 2023

In terms of individual country performance, **France** (100%) is the most mature in the quality dimension, achieving full points in all four of the underlying indicators (Figure 2). **Latvia** (95%) and **Ukraine** (94%) follow closely, both demonstrating full maturity (100%) in the ‘monitoring and

measures' indicator. Ukraine also demonstrates full maturity in the 'DCAT-AP compliance' indicator. In addition, **Denmark** (92 %) and **Poland** (90 %) are also notable performers, both achieving above 90 % maturity in this dimension. Denmark is the only country besides France to achieve 100 % maturity on the 'deployment quality and linked data' indicator, while Poland scores full points on the 'monitoring and measures' and the 'DCAT-AP compliance' indicators. Overall, 15 Member States score above the EU average of 79.7 %.

Highlight from France – automated metadata harvesting

One of the key practices highlighted in this year's report is the use of fully automated harvesting systems for metadata, whereby metadata is updated from the source rather than edited manually.

France ensures that 100 % of the essential metadata on its national portal, data.gouv.fr, is obtained automatically from the source through a comprehensive and robust system. The platform offers several methods for data publication:

- direct publication,
- publication via application programming interfaces (APIs),
- publication through harvesting.

Data.gouv.fr supports the harvesting of various metadata formats, including DCAT, Comprehensive Knowledge Archive Network (CKAN) and GeoNetwork. Data providers can set up a harvester, which the data.gouv.fr team must then validate to ensure accuracy and compatibility. This process involves establishing the mapping of fields for essential metadata while retaining additional properties for traceability.

Once harvesters are properly configured, data.gouv.fr can retrieve 100 % of the essential metadata for their model. Currently, harvested data accounts for approximately 50 % of the data.gouv.fr catalogue, highlighting the effectiveness and efficiency of this automated system in maintaining up-to-date and accurate metadata on the national portal. **Read more about this trend in Section 6.2**

Serbia (+ 17 pp), **Latvia** (+ 10 pp) and **Belgium** (+ 5 pp) demonstrated the greatest year-on-year improvement in the quality dimension. **Serbia's** improvement can be attributed to substantial progress on the 'DCAT-AP compliance indicator' (+ 50 pp). Serbia now reports that it investigates the most common causes of non-compliance with DCAT-AP standards. Additionally, 90 % of Serbia's datasets now include metadata referencing a web page where the data can be accessed. Serbia also achieved a 6 pp increase in the 'deployment quality and linked data' indicator. This increase can be attributed to the recent introduction of a model for assessing the quality of data deployment in the country.

Highlight from Serbia – discussion modules to improve metadata quality

One of the trends highlighted in this year's report is the use of data quality assessment techniques that either combine or go beyond the widely used 5-star open data model and the findable, accessible, interoperable and reusable (FAIR) principles.

In **Serbia**, each dataset comes with a discussion module through which users can share feedback regarding datasets. This feature allows users to share positive and negative opinions, report anomalies and directly suggest improvements to the dataset provider. Serbia uses this discussion module to facilitate more in-depth and actionable feedback, helping to improve the quality of data over time. This enables a rich dialogue between users and data providers, promoting continuous data quality improvement. See an [example](#) of the discussion module in action with the Address Register, a fundamental public register containing data on streets (determined by local government decisions) and house numbers across the Republic of Serbia. **Read more about this trend in Section 6.5.**

2024 quality maturity scores

Protocol order, per group of countries

- Quality dimension
- △ Metadata currency and completeness
- ◇ Monitoring and measures
- ◇ DCAT-AP compliance
- × Deployment quality and linked data

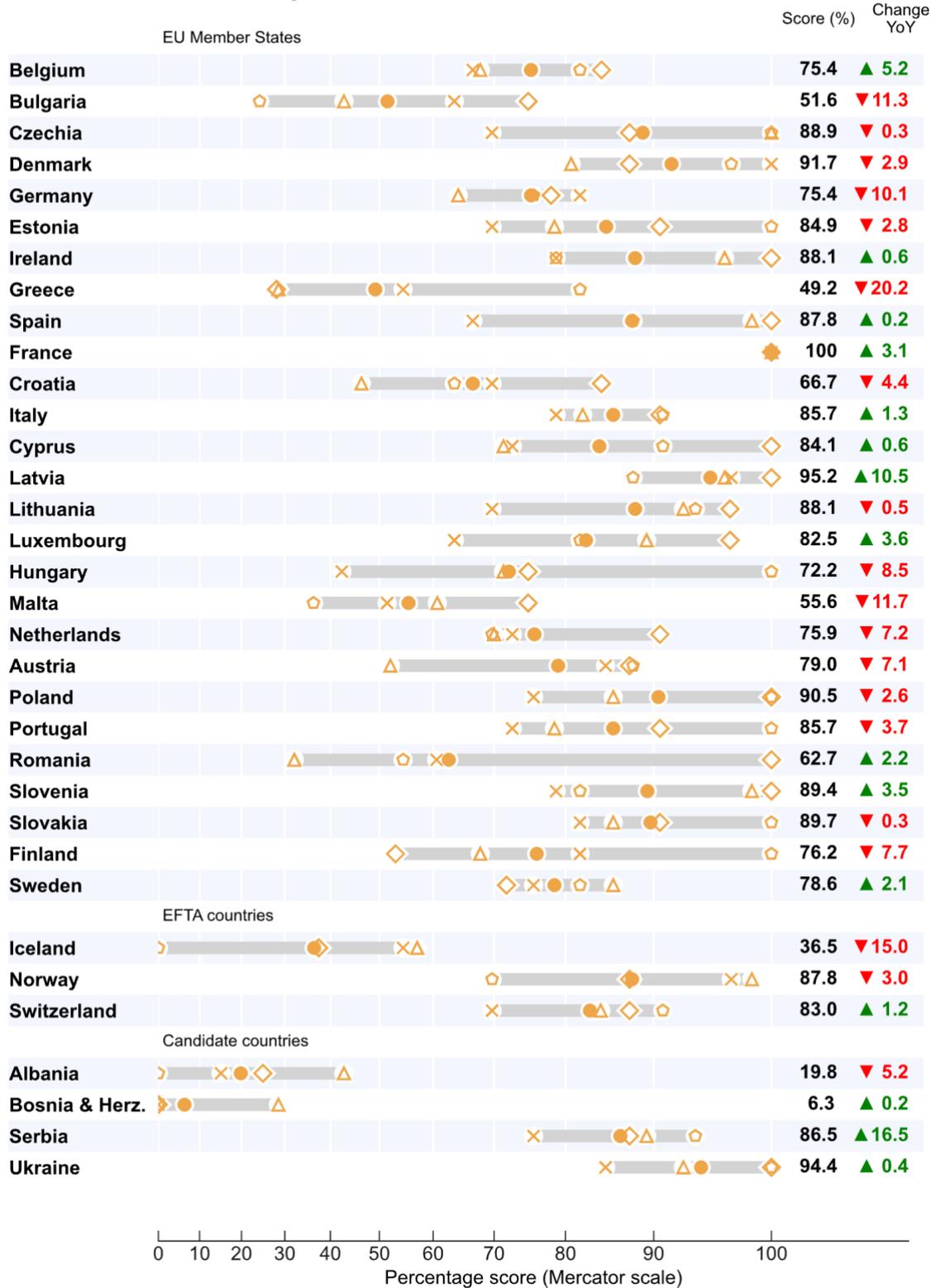


Figure 2: The scores of the majority of countries decreased on the quality dimension in 2024. (YoY: year-on-year).

Latvia's increase in its score on the quality dimension can be attributed to its 13 pp increase in the 'DCAT-AP compliance' indicator. Additionally, Latvia saw a 19 pp increase in the 'monitoring and measures indicator', which may be attributed to its recently starting to publish information on the quality of its metadata on its open data portal. **Belgium's** increase in its score on the quality dimension can be attributed to its 22 pp increase in the 'metadata currency and completeness' indicator and 14 pp increase in the 'deployment quality and linked data' indicator.

Highlight from Belgium – enhancing metadata through automated pipelines

One of the key practices highlighted in this year's report is the use of an automatic metadata harvesting process that is operated under a centralised model, whereby multiple sub portals are interconnected under a single catalogue.

In **Belgium**, using command line tools, metadata is automatically collected and then enhanced through SPARQL Protocol and RDF Query Language (SPARQL) queries and Simple Knowledge Organisation System (SKOS) data, which helps improve the structure and quality of the metadata. Once enhanced, the metadata is pushed on [GitHub](#) (a web-based platform for version control and collaboration), where different teams can access it. Finally, [data.europa.eu](#) collects the enhanced metadata from GitHub, ensuring that high-quality, standardised metadata is available at the European level. This workflow supports both automation and the enhancement of metadata quality across various systems.

Read more about this trend in Section 6.2.

6.2. Metadata currency and completeness

This indicator assesses the extent to which countries systematically ensure that their data and metadata are up to date. The indicator also investigates automatic harvesting processes, which ensure that changes at the data source are reflected with as little delay as possible on the portal where the dataset is made discoverable. Furthermore, the completeness of data that has a time component and preparations to ensure that HVDs are interoperable with other datasets on the portal are also evaluated by this indicator.

Currency of metadata

Metadata plays a crucial role in enhancing the usability and reliability of open data, and its timely update is essential for maintaining data relevance and accuracy. A predefined approach to ensuring that metadata remains up to date involves implementing systematic processes and mechanisms tailored to the specific characteristics and update frequency of different datasets. An efficient method is automatic metadata sourcing, which is when metadata is generated and updated directly from the data source without manual intervention. By adopting these practices, organisations can ensure that metadata remains accurate, relevant and aligned with the characteristics and requirements of each dataset. Table 2 presents an overview of how countries responded to the questions on this topic.

Table 2: Countries' responses to questions on the currency of metadata

| | <i>Is there a predefined approach to ensure that metadata is kept up to date?</i> | <i>What percentage of the metadata on the national portal is obtained from the source automatically rather than edited manually?</i> | <i>What is the average delay from when the metadata describing a dataset is updated at the source to when the change is visible on the portal?</i> |
|------------------|---|---|---|
| EU-27 | 26 Member States (96 %), all except Bulgaria , report having a predefined approach to ensuring that metadata is kept up to date. | Six Member States (22 %) report that 100 % of the metadata on their national portals is obtained automatically from the source. Five Member States (19 %) indicate that at least 90 % of their metadata is sourced automatically, while four Member States (15 %) report that at least 70 % of their metadata is sourced automatically. | 19 Member States (70 %) report that the average delay from when the metadata describing a dataset is updated at the source to when the change is visible on the portal is less than one day. Five Member States (19 %) indicate that this delay is typically less than one week. Croatia and Romania note that the average delay is up to one month, while Greece reports that it can extend beyond one month. |
| EFTA | All three participating EFTA countries report having a predefined approach to keeping metadata up to date. | Norway and Switzerland report that at least 90 % of the metadata on their national portals is obtained automatically from the source. Iceland reports that this figure is less than 30 % of the metadata. | Norway and Switzerland report that the average delay in updating metadata from the source is less than one day. Iceland reports that this delay can extend beyond one month. |
| Candidate | Albania , Serbia and Ukraine report having a predefined approach to keeping metadata up to date. | Ukraine reports that at least 50 % of the metadata on its national portal is obtained automatically from the source. Serbia indicates that this figure is at least 30 %, while both Bosnia and Herzegovina and Albania report that less than 30 % is sourced automatically. | Serbia and Ukraine report that the average delay in updating metadata from the source is less than one day. Bosnia and Herzegovina and Albania report that this delay can extend beyond one month. |

(Questions Q1, Q2 and Q3)

To ensure that metadata is kept up to date, most countries rely on automated harvesting systems, whereby data is continuously updated from the source. Many countries harvest metadata under a centralised model, whereby multiple sub portals are interconnected under a single (national) catalogue. This automated harvesting process comes in various forms. For example, **Estonia** and **Switzerland** report relying on a daily job scheduler, **Slovakia** conducts nightly updates via the LinkedPipes technology, and **Luxembourg** and **Slovenia** report utilising daily scripts. **Estonia, France, Luxembourg, Portugal** and **Romania** report using APIs to facilitate metadata harvesting. **Finland, France, Ireland, Luxembourg, Norway, Spain** and **Switzerland** note that their automated harvesting method accommodates different data formats (e.g. comma-separated values (CSV) and geographical JavaScript object notation (GeoJSON) and types of metadata. In addition to automated processes, some countries (e.g. **Denmark, Ireland, Austria** and **Portugal**) also note certain instances that warrant manual checks (e.g. for smaller data providers).

While most countries use time-triggered updates, such as **Italy's** weekly harvesting of the federated catalogue, **Austria** also incorporates event-triggered updates, which are activated by specific occurrences, such as legislative changes or significant weather events.

Several countries, such as **Estonia, Cyprus, Latvia** and **Slovenia**, also mention that they have legal frameworks and regulations that mandate the regular updating of metadata. In many cases, designated open data personnel and the open data portal administrators oversee the quality of metadata, conducting regular checks and informing data publishers of discrepancies. On the other hand, countries such as **Greece, Lithuania, Serbia, Slovenia, Spain** and **Switzerland** report that their metadata updates depend on the data publishers and that the data publishers set a frequency for the metadata to get updated or harvested.

Highlight from Spain – automatic metadata harvesting workflow

In **Spain**, a significant percentage (90–99 %) of the metadata on the national portal is obtained automatically from the source rather than being manually edited. This automation is facilitated through a well-structured workflow managed by the portal's back office, which includes a federation management console. This console enables publishers to manage and schedule automatic data harvesting tasks efficiently.

The process undertaken is as follows.

1. **Defining the source uniform resource identifier (URI).** Publishers begin by defining the URI of the resource description framework (RDF) / extensible markup language (XML) file that contains the source open data catalogue, ensuring that the national catalogue can accurately retrieve the metadata.
2. **Adjusting harvesting frequency.** Publishers can set the frequency for metadata harvesting based on their needs, with options including manual updates or automatic updates on a daily, weekly, biweekly or monthly basis.
3. **Automatic parsing and updating.** Once the URI and frequency are configured, the national catalogue automatically parses the RDF/XML file and updates the relevant datasets on the portal.

This automated process applies to all types of data published in source catalogues described using the RDF, ensuring that the metadata on the national portal remains consistent and up to date and requires minimal manual intervention.

Completeness of metadata

Having complete and up-to-date data is important, since the applications and reuse cases may require historical or current data to be feasible and impactful. How current this data needs to be depends on what the data is about. Datasets that represent phenomena that change in real time, such as weather or traffic data, should be updated close to real time to enable complex applications. For other datasets, a different frequency of updates may be appropriate. Gaps in a time series can also negatively affect the usability of datasets. Again, the definition of 'up to date' depends on what the data represents and the frequency with which it is collected. Table 3 presents an overview of how countries responded to the question on this topic.

Table 3: Countries' responses to the question on the completeness of metadata

| | <i>Do you undertake efforts to ensure that published data covers the full period from when it was first published?</i> |
|------------------|--|
| EU-27 | 20 Member States (74 %) report undertaking efforts to ensure that published data covers the complete time series. |
| EFTA | Iceland and Norway report undertaking efforts to ensure that published data covers the complete time series. |
| Candidate | Serbia and Ukraine report undertaking efforts to ensure that published data covers the complete time series. |

(Question Q4)

Regular monitoring, auditing and validation processes are common methods to ensure that data covers the full time series. Countries such as **Bulgaria**, **Estonia**, **Hungary** and **Serbia** report having monitoring systems in place, either through portal editors or national teams that validate the completeness of data over time. **Luxembourg** reports that it actively monitors the availability of new datasets and engages with data producers to explore the possibility of incorporating historical time-series data. Many countries also report providing direct support, guidance and recommendations to data publishers to maintain data integrity across time periods.

Furthermore, **Denmark** and **Slovakia** note that they use the attribute `dcate:temporal` within the DCAT standard and recommend it to publishers. This attribute explicitly documents the temporal scope of datasets. This is also helpful if, for example, datasets are discontinued. In these cases, the `dcate:temporal` attribute helps clarify which time periods the new datasets cover, and users can see the temporal break or shift in the data series.

Finally, automation is used in some countries to ensure the continuity of data publication and to prevent time gaps. For example, **Portugal** and **Slovakia** report that they use automated processes to update datasets and ensure no data is missed. At the same time, **Latvia's** portal includes a feature whereby data holders will see a message next to datasets if they are not updated by the specified deadline, indicating the need to update the dataset.

Highlight from Czechia – registration of data series

In **Czechia**, the national portal supports the registration of ‘data series’, which are collections of datasets connected by certain contextual relationships. These connections can be temporal (e.g. budget data over several years), spatial (e.g. lists of streets in different municipalities) or thematic (e.g. datasets from a particular system). See, for example, the data series on the [fiscal outlook](#) or the [administrative register of economic entities](#).

These measures help to ensure that datasets are comprehensive and that the data’s temporal, spatial and thematic continuity is maintained.

Interoperability of high-value datasets

The DCAT-AP annotation for HVDs can help denote HVDs, making it easier for users to identify and access them. Moreover, by adhering to this standard, national portals can ensure that their datasets are interoperable with those from other countries. Such interoperability is key to unlocking the full potential of the data, enabling more comprehensive analyses. Table 4 presents an overview of how countries responded to the questions on this topic.

Table 4: Countries’ responses to questions on the interoperability of HVDs

| | <i>Have you implemented the DCAT-AP HVDs tag to denote HVDs in your portal?</i> | <i>Besides the DCAT-AP tag, have you implemented any other measures to ensure that HVDs are interoperable with datasets from other countries?</i> |
|--------------|---|--|
| EU-27 | 17 Member States (63 %) report that they have implemented the DCAT-AP HVDs tag in their (national) open data portal(s). | 21 Member States (78 %) report that they have implemented other measures to ensure that HVDs are interoperable with datasets from other countries. |

(Questions Q5 and Q6)

Non-EU countries were not surveyed on this question, since [Commission Implementing Regulation \(EU\) 2023/138](#) on HVDs applies only to EU Member States.

Over half of Member States have implemented the DCAT-AP HVD annotations in their open data portals. However, some countries report challenges regarding compliance across all public bodies and adapting their CKAN systems to implement requirements for HVDs fully. Some countries, such as **Belgium** and **Germany**, emphasise the integration of DCAT-AP HVDs with existing geoportals. In **Germany**, Geoportal Berlin has an API connection to the Berlin open data portal, meaning that data published on the geoportal is automatically indicated in the open data portal. In **Belgium**, the properties for HVDs are collected during harvesting (i.e. the automated process of gathering and synchronising datasets) on geoportals like MetaWal, where there is also mapping between data covered under the infrastructure for spatial information in Europe (Inspire) directive ([Directive 2007/2/EC](#)), a European directive for geospatial data, and DCAT-AP HVDs. **Slovenia** and **Finland** also emphasise that they harvest data from geoportals and align it with the HVD structure. Indeed, several countries, including **Belgium**, **Spain**, **Austria** and **Romania**, with advanced initiatives related to geospatial, environmental and earth observation data, have relied on the Inspire directive’s recommendations to ensure cross-border interoperability.

In addition to the DCAT-AP HVD annotations, Member States highlight various ways to ensure that their HVDs are interoperable with datasets of other countries. For example, **Czechia** and **France** note that they directly communicate with data producers from other countries. **Denmark** and **Ireland** note that they encourage the use of standardised licences (e.g. Creative Commons (CC)) or data formats (e.g. CSV, JavaScript object notation (JSON), XML and GeoJSON) to facilitate interoperability. **Belgium** and **Ireland** also report focusing on API development as a way to ensure data interoperability.

6.3. Monitoring and measures

This indicator assesses the extent to which mechanisms are in place to evaluate and improve metadata quality and its compliance with licensing standards. Moreover, the indicator looks at the support, guidelines and tools available to assist data publishers in publishing high-quality metadata and choosing the correct licence type for their data.

Monitoring the quality of metadata on portals

Monitoring metadata quality is important for ensuring datasets are discoverable, well documented and usable by stakeholders. From manual reviews to automated systems, countries use a range of methods to ensure compliance with standards. Dashboards and reports are effective tools for monitoring and visualising metadata quality, and providing public access to these resources enhances transparency and accountability. Table 5 presents an overview of how countries responded to the questions on this topic.

Table 5: Countries' responses to questions on monitoring the quality of metadata

| | <i>Do you monitor the quality of the metadata available on your portal?</i> | <i>Do you publish information on the quality of the metadata available on the portal?</i> |
|------------------|--|---|
| EU-27 | 26 Member States (96 %), all except Finland , report monitoring the quality of metadata available on their portals. | 22 Member States (81 %), with the recent additions of Latvia, Malta and Romania , report that they publish information on the quality of metadata available on their portals. |
| EFTA | All three participating EFTA countries report monitoring the quality of metadata available on their portals. | Norway reports that it publishes information on the quality of metadata available on its portal. |
| Candidate | Serbia and Ukraine report monitoring the quality of metadata available on their portals. | Ukraine reports that it publishes information on the quality of metadata available on its portal. |

(Questions Q7 and Q8)

Most countries use automated systems to ensure metadata quality. For example, **Estonia, Germany, Lithuania, the Netherlands and Switzerland** report that they use automated systems for checking broken links, compliance with standards such as DCAT-AP and the completeness of data. Some countries such as **Bulgaria, Latvia and Ukraine** also report performing manual reviews through administrators. Additionally, **Lithuania, Austria, Poland and Romania** report enabling users to provide feedback on metadata quality, which portal administrators then take into account.

Highlight from Portugal – metadata quality box

In **Portugal**, the national open data portal (dados.gov.pt; Figure 3) provides a dedicated ‘quality’ box within the administration area to help users improve the quality of their published dataset’s metadata. This tool offers an overview of how well the dataset’s metadata is structured, highlighting areas that could be enhanced to improve discoverability and reuse. The system automatically analyses the metadata for each dataset, assessing whether it has been correctly filled out. Based on this analysis, it suggests improvements, such as adding more accurate and detailed descriptions, including additional tags, or attaching resources in more open, machine-readable formats. This proactive approach to monitoring and enhancing metadata quality ensures that contributors can easily publish high-quality, reusable data, benefiting the broader open data ecosystem.

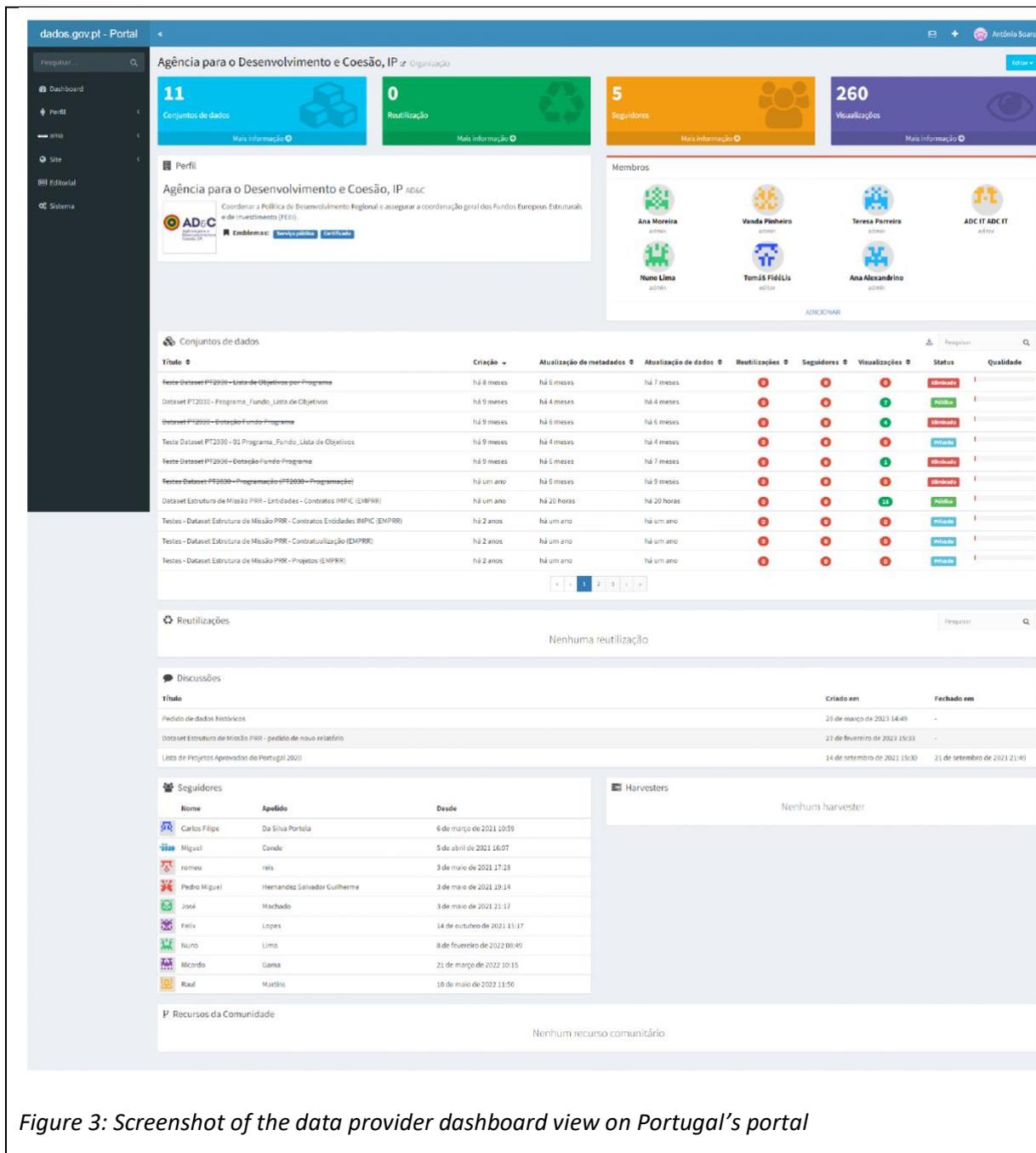


Figure 3: Screenshot of the data provider dashboard view on Portugal's portal

Many countries use dashboards and reports to monitor and visualise metadata quality. Some, such as [Italy](#), [Poland](#), [Spain](#) and [Ukraine](#), report having publicly available dashboards and reports, while others, such as [France](#), [Luxembourg](#) and [Serbia](#), report that these resources are mainly used by internal personnel. In addition, [Czechia](#) and [Germany](#) integrate their public dashboards with SPARQL (a query language for databases) to allow users to access the metadata quality data through an interface or end point. Several countries, including [Hungary](#), [Ireland](#), [Norway](#) and [Romania](#), report integrating standardised EU tools and frameworks, specifically the metadata quality assessment (MQA) methodology from data.europa.eu, to assess and monitor metadata quality.

Setting metadata standards and licensing requirements

Metadata serves as a foundational layer that describes the content, context and structure of datasets, enabling users to discover and utilise data effectively. Ensuring the quality of metadata is essential for fostering findability, interoperability and effective data sharing. Countries often set various standards and guidelines that organisations must implement to govern metadata quality and ensure the usability and reliability of open data. Licensing is a common way to govern open data and relevant metadata quality. Without a licence, data may be publicly available, but reusers will not have certainty about what permissions they have to access, use, change or share the data under copyright or database laws. Table 6 presents an overview of how countries responded to the questions on this topic.

Table 6: Countries' responses to questions on metadata standards and licensing requirements

| | <i>Do you set any standards on metadata quality that data providers must abide by?</i> | <i>Do your open data publication or licensing guidelines recommend using CC licences?</i> | <i>What percentage of the open data available on the national portal is accompanied by licensing information?</i> | <i>How many different licences are used on your portal?</i> |
|------------------|---|---|---|---|
| EU-27 | 26 Member States (96 %), all except Greece , report that they set standards on metadata quality that data providers must abide by. | 25 Member States (93 %), all except Greece and Hungary , report that their publication or licensing guidelines provide recommendations for using CC licences. Italy and Finland are the most recent additions to this group. | 21 Member States (78 %) report that more than 90 % of their datasets have licensing information. | 16 Member States (59 %) report having one to four licences on their portal. Only Belgium, Czechia and Sweden report having more than 10 licences on their portals. |
| EFTA | Switzerland reports that it sets standards on metadata quality that data providers must abide by. | All three participating EFTA countries report that their publication or licensing guidelines provide recommendations for using CC licences. | Norway and Switzerland report that over 90 % of their datasets have licensing information. | All three participating EFTA countries report having one to four licences on their portal. |
| Candidate | Serbia and Ukraine report that they set standards on metadata quality that data providers must abide by. | Albania (the most recent addition), Serbia and Ukraine report that their publication or licensing guidelines provide recommendations for using CC licences. | Serbia and Ukraine report that more than 90 % of their datasets have licensing information. | Serbia and Ukraine report having one to four licences on their portal. |

(Questions Q10, Q11, Q12 and Q13)

Countries often set mandatory metadata fields that must be filled in when publishing datasets. Specifically, many countries tend to implement variations of the DCAT-AP metadata standard, such as specifying key metadata fields as mandatory and, in some cases, mandating the inclusion of additional metadata classes compared with those required by the general DCAT-AP framework.

Most countries also have specific requirements for and recommendations on the licences that must be applied to datasets. In many countries, the CC Attribution 4.0 International (CC BY 4.0) and CC Universal (CC0) licences are mandated by laws or national strategies for public sector data publication. For example, the **Austrian** framework for open government platforms serves as an official agreement between the federal and state levels. According to this agreement, CC BY 4.0 is mandatory for Austrian public sector bodies for the publication of open government data.

Several countries offer flexibility in terms of licensing, allowing data providers to choose from various CC licences, although CC BY 4.0 and CC0 are often highlighted as preferred options. For example, in the **Netherlands**, it is mandatory to select a licence. However, publishers are free to choose which licence from a [list](#) provided by the government. In **Cyprus**, public sector bodies must seek permission from the licensing authority to use a licence other than CC BY 4.0 or CC Attribution-ShareAlike 4.0 International (CC BY-SA 4.0).

Highlight from Cyprus –tailored DCAT-AP application

Cyprus has adopted a tailored version of the DCAT-AP framework to maintain high standards of metadata quality on its national data portal. This variation includes more mandatory classes than the standard EU framework, ensuring that metadata is more comprehensive and uniform. In addition to the required fields, optional DCAT-AP fields are also available to enhance the metadata's depth.

Two standardised usage licences were selected to further support uniformity and practicality: CC BY 4.0 and CC BY-SA 4.0. Public bodies can choose the most appropriate licence when publishing datasets, ensuring clear and consistent data usage rights.

Cyprus mandates a minimum of 15 fields for metadata, which include essential information such as:

- dataset title
- description
- topics
- licence to use
- geospatial coverage
- access uniform resource locator (URL).

[Support activities for data providers](#)

Activities to support data providers with publishing high-quality data can take many forms. Documents, tools, training and tailored guidance are common methods that countries use to ensure publishers supply high-quality datasets. Table 7 presents an overview of how countries responded to the questions on this topic.

Table 7: Countries' responses to questions on support for data providers

| | <i>Do you publish guidelines and have tools to assist publishers in publishing high-quality metadata?</i> | <i>Besides providing guidelines, are regular activities conducted or mechanisms in place to assist publishers in supplying high-quality datasets?</i> |
|------------------|--|---|
| EU-27 | 23 Member States (85 %), all except Bulgaria, Greece, Croatia and Malta , report that they publish guidelines and have tools in place to assist publishers in publishing high-quality metadata. | 22 Member States (81 %) report that they conduct regular activities or have mechanisms in place to assist publishers in supplying high-quality datasets. |
| EFTA | Norway and Switzerland report that they publish guidelines and have tools in place to assist publishers in publishing high-quality metadata. | Norway and Switzerland report that they conduct regular activities or have mechanisms in place to assist publishers in supplying high-quality datasets. |
| Candidate | Serbia and Ukraine report that they publish guidelines and have tools to assist publishers in publishing high-quality metadata. | Albania, Serbia and Ukraine report that they conduct regular activities or have mechanisms in place to assist publishers in supplying high-quality datasets. |

(Questions Q9 and Q14)

Most countries publish manuals and handbooks that include information on publishing high-quality metadata. **Czechia, Denmark, Luxembourg, Hungary and Sweden** also report that they provide manuals on how to properly license metadata, helping to clarify how data can be used and whether it can be shared or modified.

Highlight from Spain – guides for improving data and metadata quality

In **Spain**, there are several guides aimed at improving the quality of both the metadata and the data itself. These include:

- [Practical Manual to Improve Open Data Quality](#),
- [How to develop a plan of measures to promote openness and reuse of open data](#),
- [Practical Guide for Publishing Linked Data in RDF](#),
- [Practical Guide for Publishing Spatial Data](#),
- [Practical Guide for Publishing Tabular Data in CSV Format](#),
- [Practical Guide for Publishing Data through APIs](#),
- [Introduction to Data Anonymisation: Techniques and case studies](#),
- [How to Implement Linked Data: Real case of the Aragón open data portal](#),
- [Open Data in Real Time: Use cases for smart cities](#),
- [Guide for priority datasets to be published by municipalities](#),
- [Data Visualisation Guide for Local Authorities](#).

Furthermore, **Belgium, Germany, Estonia, France, Italy, Lithuania and Poland** provide metadata validation tools to ensure compliance with established standards such as DCAT-AP. For example,

Germany provides a self-service tool with detailed metadata end point testing and feedback. **Estonia's** open Rihake tool allows users to describe datasets, classifiers and services and to compile data dictionaries and business glossaries, while **Belgium** provides a lightweight validator tool that checks metadata against certain standards (via GitHub). Similarly, **Italy** reports that its national open data portal includes a metadata validator that checks the conformity of its metadata to its national metadata profile DCAT-AP_IT.

In addition to guiding documents, training is a prominent way in which governments help data providers to publish high-quality datasets. For example, **Czechia, Denmark and Romania** note that they regularly host webinars about ensuring high-quality data management. In addition, **Czechia, Estonia, Hungary, Ireland, Poland and Ukraine** report that they specifically utilise e-learning to train data providers on publishing high-quality data, which also includes information on the proper procedures for metadata.

Many countries also conduct ongoing meetings with data providers to improve data quality. These are often held regularly, allowing for a continuous dialogue and exchange of knowledge (e.g. on updates and best practices) between the open data team and data providers. **Ireland and Spain** frame these meetings as 'audits on data quality', at which data providers receive personalised evaluations of their data (and metadata) quality and discuss their results. While most countries report that these meetings with data providers are one-on-one, some countries, such as **Luxembourg, Norway and Switzerland**, have routine forums and peer exchanges with established networks and groups of public sector data providers. These sessions facilitate knowledge exchange on topics such as publishing high-quality data.

Highlight from Luxembourg – three-tier approach for ensuring high-quality datasets

Luxembourg has implemented a comprehensive, multi-tier strategy to continuously improve the quality of public sector open data.

1. **Training and capacity building.** The National Institute of Public Administration offers regular training sessions dedicated to public sector open data. These sessions are available to all public sector agents and cover the national portal and related open data topics. This ensures that public officials are well equipped with the knowledge and skills necessary to manage open data effectively.
2. **Open data representatives group.** In line with the Prime Minister's directive, public sector organisations are required to appoint official open data representatives. This group facilitates regular meetings to exchange best practices, technical support and general open data information. This network of representatives enhances collaboration and ensures that organisations meet the obligations of the open data directive ([Directive \(EU\) 2019/1024](#)) and national laws. The first meeting of this series took place in April 2024, marking the beginning of this collaborative effort.
3. **One-on-one support for data owners.** Luxembourg's open data team works closely with individual data owners, providing tailored advice and technical assistance when datasets are published. This personalised approach ensures a focus on data quality, including aspects like regular updates, resource formats and metadata completeness. Additionally, Luxembourg has implemented a link-checking programme to automatically verify the availability of datasets not hosted directly on the portal. The system runs monthly tests, and any issues (e.g. broken links) are addressed by contacting the relevant organisations. This has proven effective in maintaining data availability, as demonstrated with Luxembourg's national weather data provider, Meteolux.

6.4. DCAT-AP compliance

[DCAT](#) is a World Wide Web Consortium standard designed to facilitate interoperability between data catalogues published online. [DCAT-AP](#) is an extension to DCAT – an ‘application profile’ – developed by the European Commission to improve interoperability and foster the discoverability and reuse of open data across European catalogues. The ‘DCAT-AP compliance’ indicator assesses the extent to which metadata on national portals complies with the DCAT-AP standard for describing public sector datasets and what efforts are taken to assist data publishers in following DCAT-AP.

Creating a framework for DCAT-AP compliance

Having a standard way to describe datasets helps to ensure that data catalogues from different organisations or regions are compatible. This is why many national portals follow the DCAT-AP framework or other standards to ensure interoperability with DCAT-AP. Many countries have created national extensions of DCAT-AP to tailor the general framework to their specific needs, enhancing its relevance and functionality for their contexts. Table 8 presents an overview of how countries responded to the questions on this topic.

Table 8: Countries’ responses to questions on creating a framework for DCAT-AP compliance

| | <i>Does the national portal follow the DCAT-AP framework or, if not, are standards in place to ensure interoperability with DCAT-AP?</i> | <i>Is there a national extension of the DCAT-AP standard developed for your country?</i> |
|------------------|--|--|
| EU-27 | 24 Member States (89 %), all except Bulgaria, Greece and Malta , report that their national portals follow the DCAT-AP framework or ensure interoperability with DCAT-AP. | 15 Member States (56 %) report having a national extension of the DCAT-AP standard. |
| EFTA | Norway and Switzerland report that their national portals follow the DCAT-AP framework or ensure interoperability with DCAT-AP. | Norway and Switzerland report having a national extension of the DCAT-AP standard. |
| Candidate | Serbia and Ukraine report that their national portals follow the DCAT-AP framework or ensure interoperability with DCAT-AP. | None of the participating candidate countries reports having a national extension of the DCAT-AP standard. |

(Questions Q15 and Q19)

Several countries ensure compliance with DCAT-AP by leveraging existing technical platforms or plug-ins designed with built-in DCAT-AP support. For instance, many countries note that they use the CKAN platform, which has a plug-in that allows users to describe datasets according to DCAT-AP standards. On the other hand, **Portugal** and **Serbia** report using the Udata platform, which follows the DCAT-AP standard and has tools for mapping other frameworks and standards (e.g. CKAN and operational data store (ODS)). **Ireland** and **Norway** highlight that they have implemented tools to automatically validate metadata against DCAT-AP standards.

Highlight from Ireland – automated audit tool for DCAT-AP

In **Ireland**, the national open data portal, data.gov.ie, fully adheres to the DCAT-AP framework to ensure interoperability and consistency with international open data standards. Compliance with DCAT-AP is mandated for all data published on the platform, which is outlined in the portal's [technical framework](#) and the [publishing guidelines](#).

To ensure compliance, the portal advocates using URIs, which play a crucial role in improving data discovery and interoperability across platforms. Additionally, the portal provides an audit tool specifically designed to validate datasets against DCAT-AP standards. This tool checks for missing mandatory properties, verifies the correct use of controlled vocabularies and ensures adherence to the DCAT-AP schema.

Moreover, the portal offers a range of training resources to support data providers in creating high-quality, DCAT-AP-compliant metadata. These resources include best practices, examples and tutorials, all aimed at enhancing the quality and interoperability of published datasets. Through these measures, data.gov.ie ensures that all published data meets international standards, facilitating better data sharing and reuse.

Although not compulsory, many countries have developed national extensions of the DCAT-AP standard. These countries often emphasise that their modifications are intended to better serve the needs of their national contexts, particularly for the public sector and data communities. For example, **Czechia, Italy** and **Poland** report making such modifications to comply with their specific legal frameworks or regulations that govern data. These legal obligations often necessitate changes to metadata, vocabularies or properties to ensure compliance.

Other countries report that they have made national extensions to ensure more structured and comprehensive metadata. This often involves adding additional mandatory fields or adapting vocabularies to ensure consistency and data quality. For example, the **Netherlands** reports enriching the EU standard by enabling fewer free-answer options, which it believes allows easier verification of the metadata quality.

Compliance with the DCAT-AP specifications

DCAT-AP has various metadata properties that can be used to describe data. As a specification, DCAT-AP defines a hierarchy of properties, grouped as classes, that are mandatory, recommended or optional. Table 9 presents an overview of how countries responded to the questions on this topic.

Table 9: Countries' responses to questions on compliance with DCAT-AP specifications

| | <i>What is the percentage of metadata on your portal that is DCAT-AP compliant in terms of mandatory classes?</i> | <i>What is the percentage of metadata on your portal that uses DCAT-AP recommended classes?</i> | <i>What is the percentage of metadata on your portal that uses DCAT-AP optional classes?</i> |
|--------------|--|--|---|
| EU-27 | 23 Member States (85 %) report that more than 90 % of their portals' metadata complies with DCAT-AP's mandatory classes. | 20 Member States (74 %) report that more than 90 % of the metadata on their portals follows DCAT-AP's recommended classes. | 16 Member States (59 %) report that more than 90 % of the metadata on their portals follows DCAT-AP's optional classes. |

| | <i>What is the percentage of metadata on your portal that is DCAT-AP compliant in terms of mandatory classes?</i> | <i>What is the percentage of metadata on your portal that uses DCAT-AP recommended classes?</i> | <i>What is the percentage of metadata on your portal that uses DCAT-AP optional classes?</i> |
|------------------|---|---|--|
| EFTA | Norway and Switzerland report that more than 90 % of the metadata on their portals is compliant with DCAT-AP's mandatory classes. | Norway and Switzerland report that more than 90 % of the metadata on their portals follows DCAT-AP's recommended classes. | Switzerland reports that at least 50 % of the metadata on its portal follows DCAT-AP's optional classes. For Norway , this percentage is at least 30 % and, for Iceland , it is less than 10 %. |
| Candidate | Serbia and Ukraine report that more than 90 % of the metadata on their portals is compliant with DCAT-AP's mandatory classes. | Serbia and Ukraine report that more than 90 % of the metadata on their portals follows DCAT-AP's recommended classes. | Ukraine reports that more than 90 % of the metadata on its portal follows DCAT-AP's optional classes. |

(Questions Q16, Q17 and Q18)

Despite growing compliance with DCAT-AP, not all data providers publish data that fully aligns with the DCAT-AP standard. Investigating the common causes of non-compliance can help national portals to develop strategies to help data providers improve the quality of their metadata. Table 10 presents an overview of how countries responded to the question on this topic.

Table 10: Countries' responses to the question on non-compliance with the DCAT-AP standard

| | <i>Do you investigate the most common causes of the lack of DCAT-AP compliance?</i> |
|------------------|--|
| EU-27 | 19 Member States (70 %), with Greece and Sweden as the most recent additions, report investigating the most common causes of the lack of DCAT-AP compliance. |
| EFTA | Switzerland reports investigating the most common causes of the lack of DCAT-AP compliance. |
| Candidate | Serbia and Ukraine report investigating the most common causes of the lack of DCAT-AP compliance. Serbia newly reports this. |

(Question Q20)

The most common cause of a lack of compliance with DCAT-AP is a lack of training, awareness or expertise on the standard. In other words, some data providers are unfamiliar with the requirements of DCAT-AP and do not know how to properly manage and structure data in line with the standards.

Some countries note that some compliance issues relate to the need to map or integrate metadata from different systems or translate metadata based on a different standard from DCAT-AP. **Denmark** and **Lithuania** specifically note that there are challenges associated with mapping or converting geospatial datasets structured according to the Inspire standard to/into the DCAT-AP standard.

Other causes include the cost and complexity of updating the national platform while maintaining custom settings when the DCAT-AP specification is updated. When datasets are entered manually, metadata may be incomplete or incorrectly filled, leading to deviations from the standard.

6.5. Deployment quality and linked data

This indicator examines the extent to which countries use a model, such as the Berners-Lee [5-star open data model](#) or the [FAIR principles](#), to assess the quality of data deployment. This indicator also assesses the extent to which data is available under an open licence, in structured and machine-readable formats, with URIs and links to other data sources.

Use of models for deployment quality

A model for assessing data deployment is crucial because it enables national portal teams to judge systematically and adaptively whether a dataset is more or less likely to be reused, given the quality it offers portal users. Table 11 presents an overview of how countries responded to the question on this topic.

Table 11: Countries' responses to the question on the use of models for deployment quality

| | <i>Do you use a model to assess the quality of deployment of data in your country?</i> |
|------------------|--|
| EU-27 | 24 Member States (89 %), all except Bulgaria, Hungary and the Netherlands , report using a model to assess the quality of deployment of data. |
| EFTA | All participating EFTA countries report using a model to assess the quality of deployment of data. |
| Candidate | Serbia and Ukraine report using a model to assess the quality of deployment of data. |

(Question Q23)

The 5-star open data model is a framework designed to assess the quality and openness of data based on five progressive criteria. Each level of the model corresponds to a star, with more stars indicating higher levels of openness and usability. The 5-star open data model is the most frequently cited model used by countries for assessing data quality. Some countries, such as **Cyprus** and **Ukraine**, even report that using the 5-star model is written into their national guidelines and policies.

Many countries also integrate the FAIR principles into their data quality assessments. The [FAIR](#) data principles state that it should be possible to find data, there should be information about how to gain access to the data, the data should be compatible with other data and it should be possible to reuse the data. Countries including **Belgium, Estonia, Spain, Italy, Luxembourg** and **Finland** report incorporating both the 5-star model and the FAIR principles into their models for assessing the quality of deployment of data.

Some countries utilise different assessment techniques. For example, **Denmark** has established a 'common language for data quality', which is intended as a shared reference point for discussions related to data quality issues.

Highlight from Spain – comprehensive data quality model

In **Spain**, a comprehensive approach is utilised to ensure the quality of open data deployment by leveraging both the 5-star open data model and the FAIR principles. The [5-star open data model](#) is used to classify distributions of datasets based on their publication format. In addition, the Government of Spain’s Data Office uses the FAIR principles as a basis for defining the guiding principles for the data and its infrastructures and facilitating its reuse. Furthermore, the national portal promotes the publication of open data that achieves at least the 3-star level on the scale (see an [example](#) of this promotion, which describes the process of transforming tabular datasets in CSV (3 stars) into linked and semantically enriched data (5 stars)).

Activities for data providers to ensure high-quality data

The quality of data on national portals depends on the quality of data supplied by data providers. Therefore, assisting data providers with skills and tools is one way to improve the quality of published data. Table 12 presents an overview of how countries responded to the question on this topic.

Table 12: Countries’ responses to the question on activities for data providers to ensure high-quality data

| | <i>Do you conduct activities to promote and familiarise data providers with ways to ensure higher quality data?</i> |
|------------------|--|
| EU-27 | 26 Member States (96 %), all except Bulgaria , report conducting activities to promote and familiarise data providers with ways to ensure higher quality data. |
| EFTA | All participating EFTA countries report conducting activities to promote and familiarise data providers with ways to ensure higher quality data |
| Candidate | Albania, Serbia and Ukraine report conducting activities to promote and familiarise data providers with ways to ensure higher quality data. Albania is the most recent addition to this group. |

(Question Q24)

Many countries use training programmes and workshops to educate data providers on best practices for data quality. This includes training on the 5-star open data model and FAIR principles. **Estonia** and **Ukraine** note that their general training schemes include theoretical knowledge and practical skills for effectively publishing datasets. **Norway** and [Spain](#) also report publishing blogs for data providers that disseminate best practices. Several countries have developed guidelines and best practices to assist data providers in understanding and achieving high data quality.

Additionally, many countries are promoting efforts to engage with data providers more, providing feedback on their datasets and sharing best practices. For example, **Belgium, Greece and Poland** report having regular consultations with data providers, working one-on-one to improve their data quality.

Highlight from Sweden – the data ambassador programme

In **Sweden**, the [data ambassador programme](#), launched by the Swedish Agency for Digital Government, is a pioneering educational initiative aimed at enhancing the understanding and implementation of open data practices among public sector employees. Developed in response to the increasing need for effective data sharing following the enactment of the Open Data Act in August 2022, the programme targets individuals working operationally with open data and data sharing within public organisations.

This digital learning initiative offers a self-paced format that includes recorded videos, references for further reading and short knowledge assessments, focusing on fundamental processes and concepts of open data in alignment with the Open Data Act. By equipping participants with essential knowledge on how to share and utilise data effectively, the programme fosters an environment in which data sharing becomes a strategic resource.

The agency plans to expand the programme by offering additional courses tailored to various stakeholders, including subcontractors, managers and legal professionals. This initiative aims to strengthen the understanding of data as a strategic resource, ultimately promoting higher data quality across the public sector.

Deployment quality

Several best practices can enhance the accessibility and reusability of open data. These include ensuring datasets are made available under an open data licence (e.g. CC) and having licences provided in a structured format. Additionally, it is good practice to ensure that datasets are in an open and machine-readable format (e.g. CSV, JSON and XML) and to assign URIs to the datasets. Finally, datasets should also be linked to various sources, which through the use of URIs can expand the dataset’s context and relevance. Table 13 presents an overview of how countries responded to the questions on this topic.

Table 13: Countries’ responses to questions on deployment quality

| | <i>What percentage of datasets are made available under an open licence?</i> | <i>What percentage of licences are provided in a structured data format?</i> | <i>What percentage of datasets are provided in an open and machine-readable format?</i> | <i>What percentage of datasets use URIs?</i> | <i>What percentage of datasets link to other sources?</i> |
|--------------|---|---|---|--|---|
| EU-27 | 22 Member States (81 %) report that over 90 % of their datasets have an open licence. | 19 Member States (70 %) report that over 90 % of their datasets have structured licence data. | 17 Member States (63 %) report that over 90 % of their datasets are in a machine-readable format. | 8 Member States (30 %) report that over 90 % of their datasets use URIs. | 5 Member States (19 %) report that over 90 % of their datasets are linked to other sources. |

| | <i>What percentage of datasets are made available under an open licence?</i> | <i>What percentage of licences are provided in a structured data format?</i> | <i>What percentage of datasets are provided in an open and machine-readable format?</i> | <i>What percentage of datasets use URIs?</i> | <i>What percentage of datasets link to other sources?</i> |
|------------------|--|--|--|---|--|
| EFTA | Switzerland reports that over 90 % of its datasets have an open licence. | All three participating EFTA countries report that over 90 % of their datasets have structured licence data. | Iceland reports that over 90 % of its datasets are in a machine-readable format. | Norway reports that over 90 % of its datasets use URIs. | Norway reports that over 90 % of its datasets are linked to other sources. |
| Candidate | Serbia and Ukraine report that over 90 % of their datasets have an open licence. | Serbia reports that over 90 % of its datasets have structured licence data. | Serbia and Ukraine report that over 90 % of their datasets are in a machine-readable format. | None of the participating candidate countries reports using URIs. | None of the participating candidate countries reports linking datasets to other sources. |

(Questions Q25, Q26, Q27, Q28 and Q29)

6.6. Pilot indicator: automated tests of metadata quality

Pilot indicator – automated tests

In addition to gathering qualitative data about metadata quality, such quality can also be quantitatively assessed. The [MQA](#) is a tool designed to evaluate the quality of metadata harvested by data.europa.eu. It enables data providers and portals to assess their metadata and receive recommendations for improvement.

The MQA's methodology (which is undergoing a recalibration) examines five specific questions, which focus on:

- compliance with DCAT-AP and related standards,
- the disclosure of information beyond DCAT-AP requirements,
- the accessibility of referenced data,
- the machine readability of data formats
- the use of licences.

As a pilot project in the ODM assessment, we analysed five indicators from the MQA, calculating summary statistics across national catalogues that were findable on data.europa.eu. Certain countries, including **Albania, Bosnia and Herzegovina, Estonia, Greece** and **Malta**, were not

assessed, as their primary open data catalogues were not findable on data.europa.eu. The results presented below are the percentages of datasets across the selected catalogues that met each criterion. The metrics were extracted from the MQA on 31 October 2024. The findings did not contribute to the countries' maturity scores.

The **machine readability** indicator evaluates if a distribution is in a machine-readable format based on data.europa.eu's [GitLab repository vocabulary](#).

- 65 % of the distributions assessed are machine readable. **Bulgaria** scored 100 % on this indicator.

The **DCAT-AP compliance** indicator evaluates metadata conformity with the DCAT-AP standard using the shapes constraint language (SHACL) validation from data.europa.eu. SHACL is a recommendation from the World Wide Web Consortium and is used for validating RDF graphs against a set of shapes.

- 21 % of the distributions assessed are DCAT-AP compliant, with **Hungary** scoring 100 % on this indicator.

The **download URL** indicator evaluates whether direct access to data is provided via a download URL.

- 35 % of the distributions assessed included a direct link, with **France, Italy, Cyprus, Lithuania, Poland, Slovakia** and **Finland** scoring 100 % on this indicator.

The **licence information** indicator evaluates if distributions specify licence details, facilitating reuse.

- 55 % of the distributions assessed provide licence information, with **Cyprus, Czechia, Germany, Iceland, Italy, Luxembourg, the Netherlands, Spain** and **Switzerland** achieving 100 % on this indicator.

The **licence vocabulary** indicator evaluates the accuracy of licence specifications (e.g. correctly versioned CC licences). The specifications are derived from the [FAIR principles](#). The MQA recommends and credits the usage of controlled vocabularies. The data.europa.eu portal publishes its controlled vocabularies on [GitLab](#). The vocabularies are derived from the [EU vocabularies](#).

- 46 % of the distributions assessed include licence information that matches controlled vocabularies, with **Bulgaria, Ireland, Cyprus** and **Portugal** scoring 100 % on this indicator.



2024 Open Data Maturity Report
Celebrating 10 years of ODM