WEBINAR

Open Data Maturity 2022: Diving deeper into the quality dimension



18 April 2023

14.00 — 16.00 CET

Rules of the game



The webinar will be recorded



Please mute yourselves during the webinar



Please reserve 3 min after the webinar to help us improve by filling in our feedback form



For questions, please use the ClickMeeting chat.



Introduction



Simon Steuer
Publications Office of the EU
Head of Sector



Giulia Carsaniga Research and Policy Lead Consultant, Capgemini



Aleš Veršic
Government Office for Digital
Transformation,
Head of Division for the data
economy



Mykhailo Kornieiev
Head of open data
development –
Ministry of Digital
Transformation of Ukraine



Yanina Liubyva
State expert on open data
development - Ministry of Digital
Transformation of Ukraine



Agenda

14.05 – 14.25	Open Data Maturity 2022 with a deep dive into the quality dimension – <i>Giulia Carsaniga</i>
14.25 – 14.50	Open data best practices from Slovenia – Aleš Veršic
14.50 – 15.15	Open data best practices from Ukraine - Mykhailo Kornieiev & Yanina Liubyva
15.15 – 15.45	Q&A
15.45 – 15.50	Recap and closing



Diving deeper into the quality dimension



Introducing the Open Data Maturity assessment

What

■ 8th consecutive assessment of the state of open data in Europe, with 35 participating countries



















Why

Gain insights in the state of open data and challenge countries to improve

How

Self-assessment: national open data teams fill out a questionnaire on 4 dimension of maturity





Introducing the Open Data Maturity dimensions and

indicators



- Policy framework
- Governance
- Implementation



- Strategic awareness
- Measuring re-use
- Creating impact
 Governmental, Social, Environment & Economic



- Portal features
- Portal usage
- Dataprovision
- Portal sustainability



- Currency and completeness
- Monitoring and measures
- DCAT-AP compliance
- Deployment quality and linked data

2022 Revised Approach

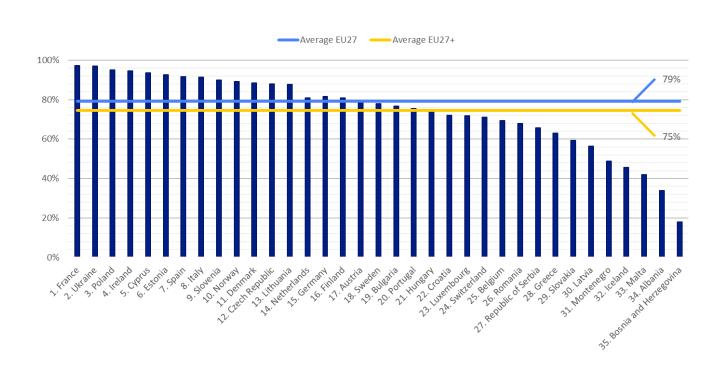
- **Streamline of questions** to better include regional and local realities and EU priorities
- More focus on high-value datasets to test countries' level of preparedness ahead of the regulation
- Re-structuring of the Open Data impact dimension to:
 - better acknowledge the chall enge of assessing impact
 - better distinguish between re use and how impact is created

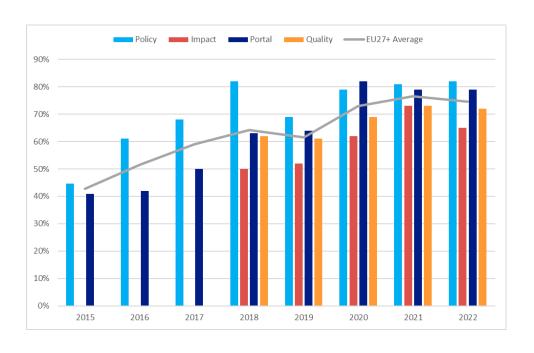
These four dimensions allow a comprehensive assessment of open data developments across Europe



Results 2022 (EU27+)

- In 2022, the average open data maturity score of the EU27+ countries is 75% vs. 78% (2021).
- Policy is the most mature dimension (82%), followed by portal (79%), quality (72%), and impact (65%).
- France (97,2%) and Ukraine (96,9%) lead the way of open data developments in Europe.

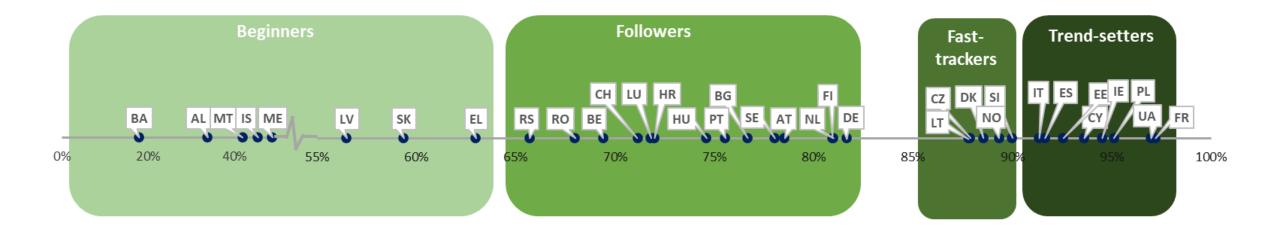






Countries clustering 2022 (EU27+)

- The maturity of European countries is concentrated in the higher end of the spectrum (above 65%).
- The **fast-trackers show highly similar scores**, as the cluster is concentrated on a range of 3% (88 to 91 percentage points).
- Remarkable is the journey of
 - **Belgium** from beginners in 2021 to followers in 2022;
 - Czech Republic from followers in 2021 to fast-trackers in 2022;
 - Italy and Cyprus from fast-trackers in 2021 to trend-setters in 2022.





The quality dimension and its indicators

Dimension 4: Open Data Quality 1. Currency (up to date) and completeness A pre-defined approach is in place to ensure metadata is up to date. Mechanisms are in place to ensure that changes at the source are reflected with the least amount of delay on the national portal. The portal provides access to a vast range of data, both historical and current. Preparation to ensure interoperability of high-value datasets alongside the datasets of another country has started. 2. Monitoring and measures Mechanisms are in place to monitor the quality of the metadata. Information on the quality of the metadata is available to the broader public. Guidelines and/or tools are available to assist data providers in choosing the right type of licence for their data. The compliance level in terms of correct licensing information is monitored. Support (e.g., in terms of documentation, tools, helpline etc.) is in place to assist data providers in improving the quality of data. 3. DCAT-AP Compliance Guidelines and materials to help data providers ensure compliance with DCAT-AP are available on the national portal. Compliance with the DCAT-AP standard in terms of mandatory, recommended and optional classes, is monitored. Monitoring activities of the percentage of accessible distributions (i.e., availability of AccessURL and DownloadURL) are in place. 4. Deployment quality and linked data A model (such as 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. The percentage of published open data that complies with your chosen quality model is known. The improvements in terms of quality of open data deployment are monitored.

Assessing the degree to which countries have a systematic approach to ensure that (meta)data is up to date.

Assessing the support, guidelines, and tools available to help publish high-quality metadata and select the best license type

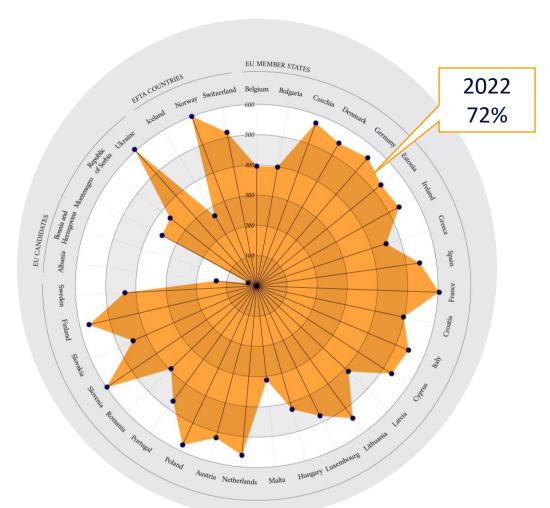
Assessing DCAT-AP compliance and the reasons for using it

Assessing the quality of deployment of data e.g. through 5-star

systems and FAIR principles



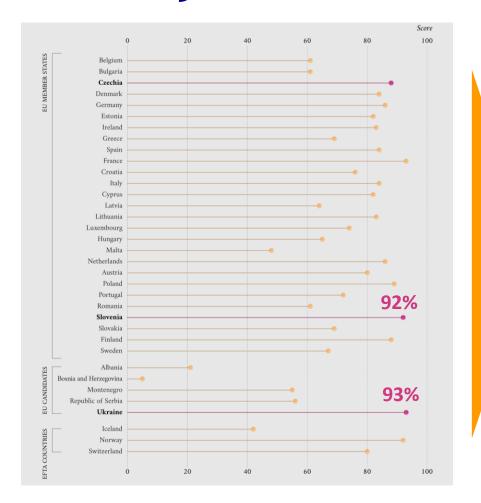
Quality results 2022 (EU27+)

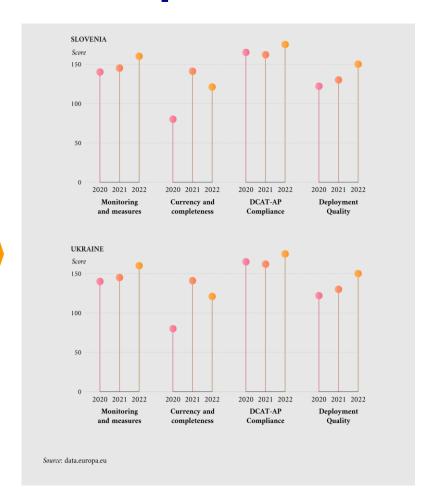


- In 2022, the **quality dimension** slightly **decreased for the EU27+:** from 73% in 2021 to 72% in 2022.
- The decrease is due to a **lower score in two indicators**:
 - Currency and completeness: 74% to 68%
 - Monitoring and measures: from 84% to 79%
- Conversely, DCAT-AP compliance increased by 6% and deployment quality by 2%.



Quality results 2022: Best practices

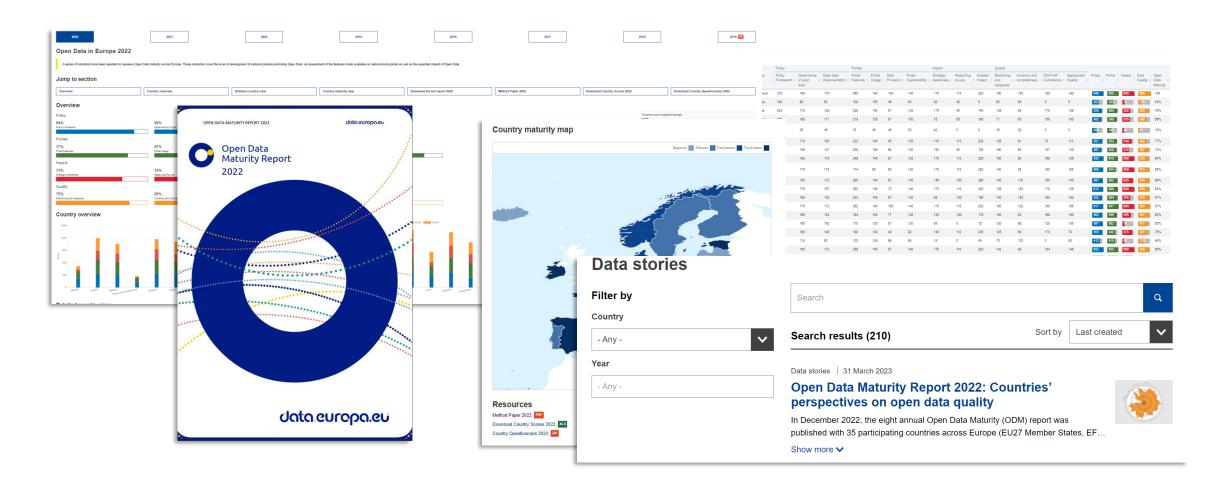




- The most mature indicator for both Slovenia and Ukraine in 2022 is monitoring and measures (100%), while the least is deployment of quality.
- They differ however in their second-best indicator: for Slovenia is currency and completeness (96%) and for Ukraine is DCAT-AP compliance (97%).



Discover all ODM-related material!





Open data best practices of Slovenia











Webinar on ODM and the quality dimension - Slovenia

Aleš Veršič, M.A.

Government Office for Digital Transformation,

Head of Division for the data economy





- Act on Access to Public Information (ZDIJZ)
- Regulation on the transfer and reuse of information of a public nature
 - Annex 1: More detailed description of metadata of public records and computerized collections
 - Annex 2: List of datasets that are high value datasets



- Annex 1: More detailed description of metadata of public records and computerized collections
- 1. Name of the institution (varchar text from code list)
- 2. The name of the dataset (free varchar text)
- 3. Theme (code list)
- 4. Content description (free varchar text)
- 5. Legal bases (Using unique ID with link to law in Legal information System)



- Annex 1: More detailed description of metadata of public records and computerized collections
- 6. Keywords (EUROVOC and GEMET descriptors) (text)
- 7. The area to which the contained data refers (geolocation)
- 8. Data availability (code list)
- 9. Any special conditions for use or reuse (code list with extra text if there is any limitation)
- 10. Possible conditions regarding the charging of the price for reuse (code list with extra varchar text or URL if there is any limitation)



- Annex 1: More detailed description of metadata of public records and computerized collections
- 11. Possible internet address for accessing or viewing the data (URL)
- 12. Possible internet address to open data (machine readable format) (URL)
- 13. Date of creation (date)
- 14. Data refresh frequency (code list based on ISO 19115)
- 15. Record format



- Annex 1: More detailed description of metadata of public records and computerized collections
- 16. Language of record (code list based on ISO 639-2)
- 17. A description of the possible connection of the data with other public records or databases or an indication that there is no such connection
- 18. Responsible person (text)
- 19. Indication that it is a dataset that belongs to the high-value datasets (yes/no)



MONITORING AND MEASURING METADATA QUALITY

Editorial process

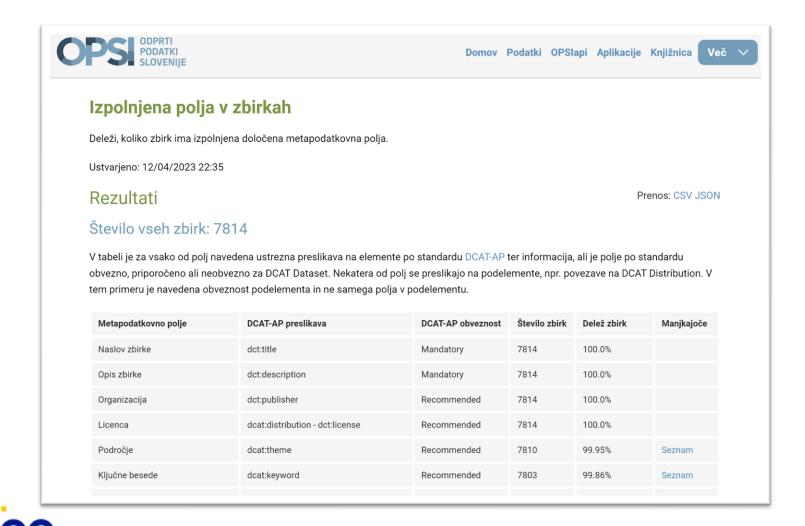
 Harvesting (from the other portal e. g. Geoportal of Slovenia, push or pull method of 3 different files (new, update, delete) with all metadata



- OPSI Portal framework: CKAN
- Plugins:
- ckanext-report
- ckanext-dcat





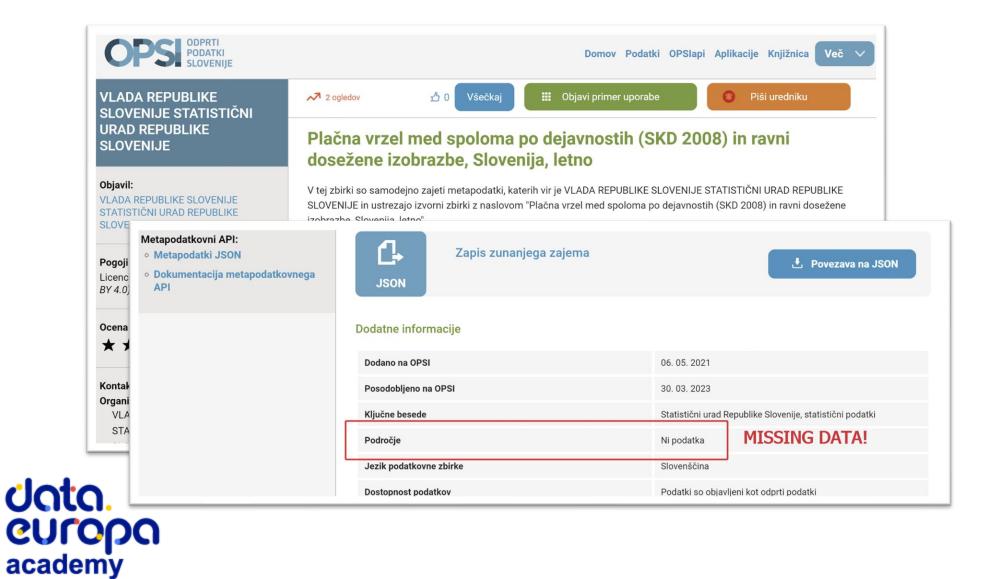


Displayed metadata	DCAT-AP mapping	Requirement level		et number coverage	Outlier list
Metapodatkovno polje	DCAT-AP preslikava	DCAT-AP obveznost	Število zbirk	Delež zbirk	Manjkajoče
Naslov zbirke	dct:title	Mandatory	7814	100.0%	
Opis zbirke	dct:description	Mandatory	7814	100.0%	
Organizacija	dct:publisher	Recommended	7814	100.0%	
Licenca	dcat:distribution - dct:license	Recommended	7814	100.0%	
Področje	dcat:theme	Recommended	7810	99.95%	Seznam
Ključne besede	dcat:keyword	Recommended	7803	99.86%	Seznam



```
"results:": [
    "https://podatki.gov.si/dataset/surs0782530s",
    "https://podatki.gov.si/dataset/surs0782525s",
    "https://podatki.gov.si/dataset/surs0782535s",
    "https://podatki.gov.si/dataset/surs0782540s"
]
}
```





DCAT-AP field	Requirement level	Completeness
dct:title	Mandatory	100.0%
dct:description	Mandatory	100.0%
dct:publisher	Recommended	100.0%
dcat:distribution - dct:license	Recommended	100.0%
dcat:theme	Recommended	99.95%
dcat:keyword	Recommended	99.86%
dcat:contactPoint - vcard:fn	Recommended	98.35%
dcat:distribution - dcat:accessURL	Recommended	93.43%
dct:publisher - foaf:name	Recommended	92.44%
dcat:distribution - dcat:downloadURL	Recommended	86.43%
dct:issued	Optional	100.0%
dct:modified	Optional	100.0%



DCAT-AP field	Requirement level	Completeness
dct:accessRights	Optional	100.0%
dct:language	Optional	99.99%
dct:spatial	Optional	99.72%
dcat:contactPoint - vcard:hasEmail	Optional	97.82%
foaf:page	Optional	89.71%
dct:conformsTo	Optional	83.08%
dct:accrualPeriodicity	Optional	79.92%
dct:relation	Optional	36.59%
dct:temporal	Optional	31.97%
dcat:contactPoint - vcard:hasTelephone	Optional	28.69%
dct:temporal	Optional	21.21%
owl:versionInfo	Optional	7.52%



RECOMMENDATIONS OR LEARNINGS REGARDING PUBLISHING REAL-TIME DATA

- Definition of real-time
- Challenges on distribution side
 - DDoS attack (limited access via API manager)
 - Network
 - Availability and performance of the data (up time of the service)
- Security issue
- Amount of data
- Data formats, schemas especially for the same topic from different sources (data space)
- Data Quality (pre-processing or raw data)
- Time sync from different sources (e.g. weather data 7:00, 7:05, time zone, ...)



NEXT STEPS OR THINGS REGARDING OPEN DATA QUALITY

- Challenges in harvesting vs manual publication
 - Missing fields in harvesting (reject all datasets?)
 - Multilingual issue
 - Manual have more control, but still depend on human error
 - Mapping metadata schema with harvesting source (using the same code list if e.g. name of institutions. Have a catalog of existing code lists)
- Contact data is depending on internal distribution process (email vs phone)
- DQ of metadata vs. dataset (should every dataset have a QA/QC report?)



Thank you





Aleš Veršič aversic@gov.si







Open data best practices of Ukraine

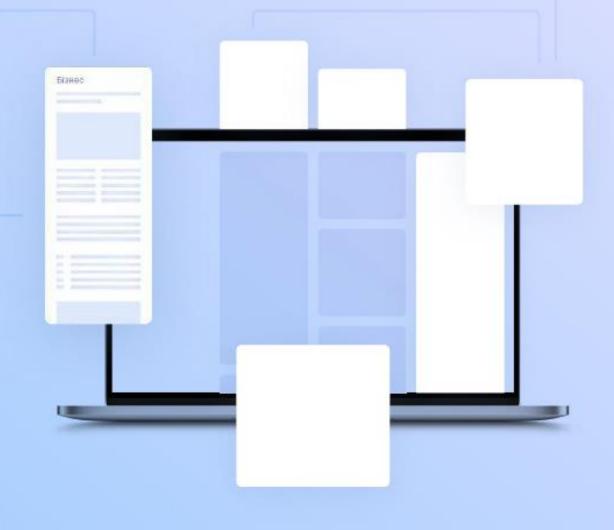






Єдиний державний веб-портал відкритих даних

Open Data Maturity 2022: Diving deeper into the quality dimension



Open Data policy in Ukraine:



Legislation and comprehensive approach to the open data and access to public information:

- Law of Ukraine "On Access to Public Information"
- Resolution of the Cabinet of Ministers of Ukraine "On approval of the regulation on data sets subject to publication in the form of open data"
- Scorecard for publishing and updating open data on the Unified State Open Data Web Portal
- Principles of International Open Data Charter



Ukrainian approach to improving the quality of open data:

- Continuous improvement of knowledge and skills of data providers in working with open data
- Developing regulation of data formats and requirements for publication of datasets
- Annual evaluation of publication and update of open data on the portal
- Development standards (requirements) for the publication of the most common datasets and of high value datasets
- "Moderation" mechanism on National Open Data portal

Monitoring the quality of datasets

The monitoring quality of datasets published by data providers on the portal consists of several levels of verification, monitoring and re-monitoring.

- 1. Before all datasets are published, they are moderated the authorized person checks the correct completion of the "dataset's passport" and metadata compliance with approved standards.
- 2. In the second stage, the portal automatically checks the data set files for logical /syntactic/semantic errors (Applies to files published in .xls(x), .xml, .json formats)

Only if two checks (manual and automatic) successfully passed dataset is published on the portal.

3. The Ministry shall continuously monitor already published datasets for their timely updating and updating in accordance with the approved requirements (standards). This applies to both the dataset's metadata and the files that contain the information.

Monitoring and measures (DCAT-AP Compliance)

Mandatory fields of dataset's metada:

identification number of the dataset;

dataset name (up to 254 characters);

a brief description of the content of the dataset (up to 4000 characters) indicating the information that is part of the dataset, but is not made public due to limited access to it (indicating the type of information with limited access for each element);

the basis and purpose of collecting the information contained in the dataset;

information about the language of the information contained in the dataset:

the format(s) in which the dataset is available;

dataset compression format (if such compression is available);

the date and time of the first publication of the dataset;

date and time of last updates to the dataset;

the date of relevance of the data in the dataset;

the frequency of updating the dataset;

keywords that reflect the main content of the dataset;

a hyperlink to the data set (downloadable electronic file or application programming interface);

a hyperlink to the dataset structure (electronic download file or application programming interface);

information about the data provider who owns the data set;

Monitoring and measures (DCAT-AP Compliance)

Mandatory fields of dataset's metada:

identification code of the manager of information in the Unified State Register of Legal Entities, Individual **Entrepreneurs and Public Organizations;**

code according to KOATUU of the territory over which the jurisdiction of the information manager extends;

information about the responsible person of the data provider, who is responsible for publishing information in accordance with the Law of Ukraine "On Access to Public Information@ and his e-mail address;

hyperlinks to previous versions of the dataset;

dataset version number.



Datasets Organizations Analytics FAQ Diia.Open data

Additional Info

Field	Value
Maintainer	Любинський Микита Владиславович
Version	1.0
Last Updated	April 11, 2023, 7:11 PM (UTC+03:00)
Created	April 11, 2023, 3:19 PM (UTC+03:00)
Update Frequency	Once a quarter
EDRPOU code	37472062
KOATUU identifier	8039100000
Purpose of collecting information	Закон України Про внесення змін до деяких законодавчих актів щодо удосконалення порядку надання адміністративних послуг у сфері будівництва та створення Єдиної державної електронної системи у сфері будівництва
Dataset ID	5255ea7e-6f94-41e4-86eb-ad578aea853c
Passport	Download
Datapackage	Download

n accordance with the Law of Ukraine "On Access to Public Information", public information in the form of open data is made public for free and free of-charge access to it. Open data shall be allowed for free re-use and dissemination.

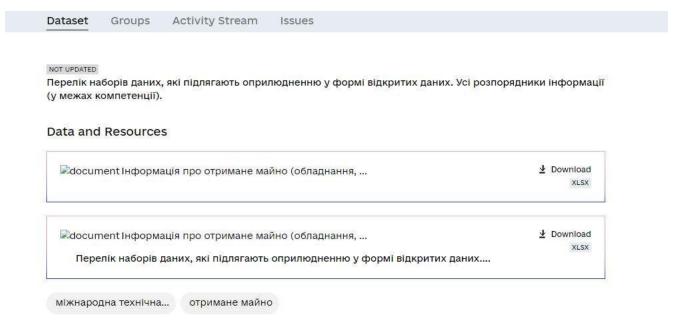
Any person may freely copy, publish, disseminate, use, including for commercial purposes, public information in the form of open data, in combination with other information or by including in their own product, with the mandatory reference to the source from which such information was received.

The mandatory reference to the source from which such information was received (including the hyperlink to the web page of the publisher's open data) is the condition for any further use of open data.

Currency and completeness

Frequency of datasets and it's metadata updates defined by Ukrainian legislation approximately for each type of dataset. The portal automatically checks if the metadata is updated in time (acording to the chosed frequency), if the datasets and it's metadata isn't updated in-time, the portal sends to the data provider a reminder to update the dataset and all neseccery metadata. The non-updated datasets is marked by the portal automatically.

If the dataset (or metadata) is not updated in time, the portal sends a corresponding message to the data provider and highlights the dataset as not updated in time.



Annual evaluation of publication and update of open data on the portal

Data providers must annually provide the Ministry of Digital Transformation with the assessment of the publication and updating open datasets on the national portal.

The assessment consists of five dimensions:

- open data policy;
- the potential of open data processing and analysis;
- quality of open data;
- user engagement;
- the impact of open data.

In the quality dimension, data providers evaluate each dataset separately according to various criteria, primarily for compliance with the principles of the international open data, machine readability, and other requirements.

Improving data providers skills

Face-to-face training (both offline and online) for representatives of central and local authorities.

The trainings concern the main areas of development of open data sphere:

- setting up the implementation of the open data policy,
- developing the regulatory and legal framework (internal),
- conducting an audit of data sets, and improving their quality by using the standards for data publication.

Publication of high-value datasets

Legal component

assessment of the dataset and the completeness of the information it carries, in accordance with the legislation and the principles of the International Open Data Charter.

Technical component

creation of recommendations for the use of correct formats for files, as well as recommendations for improving existing information systems in the direction of automatic publication of datasets according to defined standards.



Next plans to improve open data quality

Approving the DCAT-AP 2.0 specification at the level of a resolution of the Cabinet of Ministers of Ukraine (recommended and optional properties)

expand the functionality of the portal so that datasets have more detailed metadata, which greatly facilitates the use and reuse of datasets and increases the impact of each individual data set.

Automation of the process of datasets moderation. (Now it works in semi-automatic mode and needs direct control from the portal team.

Implementation into domestic legislation the European Commission IMPLEMENTING REGULATION on laying down a list of specific high-value datasets and the arrangements for their publication and re-use (C/2022/9562)

Development of standards for the publication of high-value datasets according to the criteria and recommendations of the European Commission.

Contact us



Yanina Liubyva

State expert on Open Data development -**Ministry of Digital Transformation of Ukraine**





Mykhailo Kornieiev

Head on Open Data development -**Ministry of Digital Transformation of Ukraine**

Questions & Answers





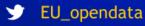
Stay up-to-date on our 2023 activities!

- Webinar series dedicated to use cases continued – 25 April, 'Stories of Use Cases: Open Data to foster European's economy and way of life'
- New webinar series on data spaces and open data – to start on 12 May
- Webinar series dedicated to the Open
 Data Maturity dimensions To be
 continued during July



Sign up for the newsletter: data.europa.eu/newsletter

Follow us on social media:



- in Publications Office of the European Union
- f data.europa.eu







Please provide your feedback!







Thank you

data. europa academy

