

data.europa.eu and the European common data spaces

Elena Simperl

13-12-2022





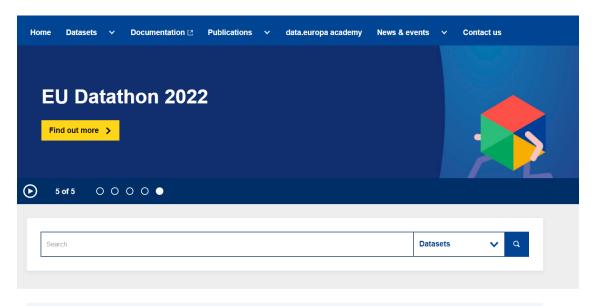




data.europa.eu







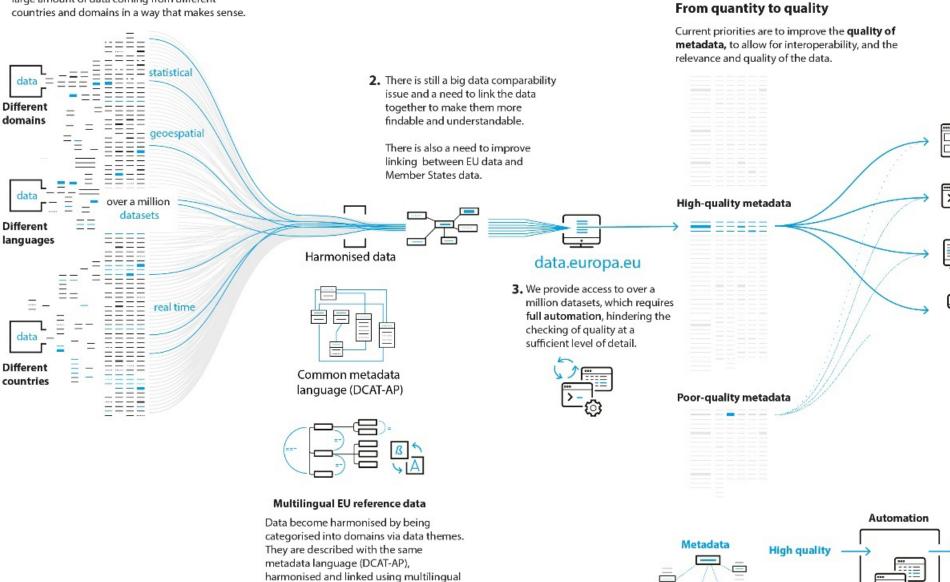


Events



Unique features of data.europa.eu

1. The challenge is to present a heterogeneous and large amount of data coming from different countries and domains in a way that makes sense.



Poor quality

EU reference data.



Institutions,

goverments,

Professionals,

developers,

journalists,

research community, etc.

Systems,

High-quality

service

service

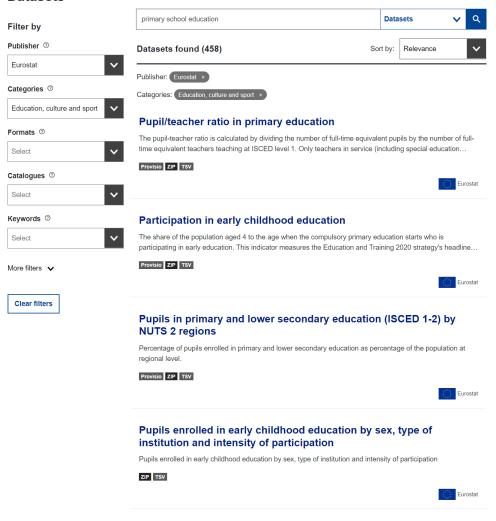
Poor-quality

software, etc.

companies, etc.

agencies,

Datasets







Show more >

Identifiers

Catalogue



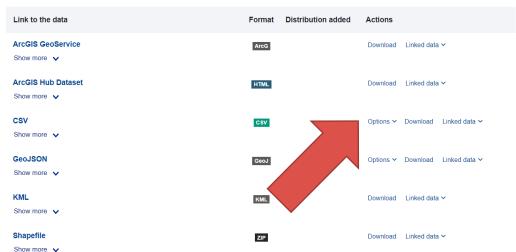
The title and description of this dataset are machine translated.

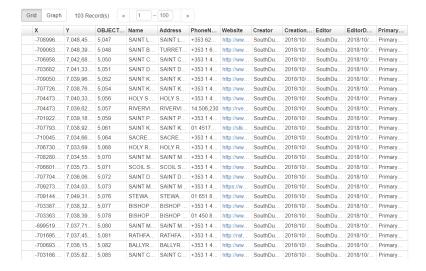
1bdcd507-61b6-40f3-9ea0-352c7991bb7c

Keywords (7)



Distributions (6)









Overview

Cybersecurity

(ÉU)

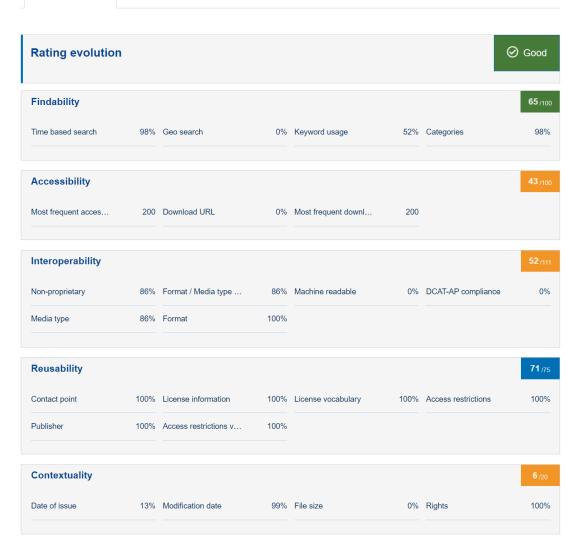
Weight **Top 12 Catalogues** Top 20 all Catalogues 100 Points 100 Points 110 Points 75 Points 20 Points 405 Points Country Name Findability Accessibility Interoperability Reusability Contextuality Rating European Union 100 / 100 100 / 100 80 / 110 75 / 75 5/20 Excellent Intellectual Property Office (EU) 98 / 100 97 / 100 77 / 110 75 / 75 10 / 20 European Excellent Central Bank (EU) 100 / 100 100 / 100 60 / 110 Directorate-75 / 75 15 / 20 Good General for Migration and Home Affairs (EUROPE) Directorate-90 / 100 100 / 100 72 / 110 75 / 75 10 / 20 Good General for Education, Youth, Sport and Culture (EUROPE) European 100 / 100 100 / 100 60 / 110 75 / 75 10 / 20 Good Research Council Executive Agency (EU) 100 / 100 100 / 100 60 / 110 75 / 75 Executive 10 / 20 Good Agency for Small and Medium-sized Enterprises (EUROPE) 90 / 100 100 / 100 70 / 110 75 / 75 10 / 20 European Good Agency for Safety and Health at Work (EU) 80 / 100 100 / 100 80 / 110 75 / 75 European 5/20 Good Political Strategy Centre (EU) European Union 90 / 100 100 / 100 60 / 110 75 / 75 13 / 20 Good Agency for

Eurostat



 Catalogue dashboard
 Distribution accessibility issues (427)
 DCAT-AP Schema Violations (57 092)





 Datasets by data theme
 Datasets per origin
 Datasets per origin and catalogue

Home > Statistics > Historical Catalogue Statistics > Datasets per origin

Jata academy 💐

Total Number of Datasets: 1 277 637

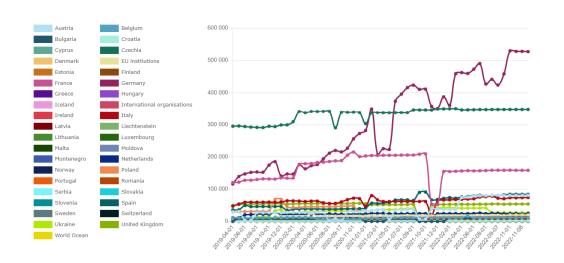
	Data theme ^	Trend per year	Percentage of total datasets	Number of Datasets
\$	Agriculture, fisheries, forestry and food	-0.00%	27.75 %	355 802
⊕ g	Economy and finance	0.00%	5.5 %	69 990
	Education, culture and sport	0.00%	2 %	24 508
Ô	Energy	0.00%	0.75 %	8 030
*	Environment	0.00%	21.5 %	275 786
	Government and public sector	-0.00%	7.25 %	94 186
8	Health	0.00%	1.5 %	19 378
Hr.	International issues	0.00%	0.09 %	1 138
Ą.	Justice, legal system and public safety	0.00%	14.75 %	188 985
92	Population and society	-0.00%	4.25 %	55 873
(4)	Regions and cities	-0.00%	3.75 %	46 434
48	Science and technology	0.00%	7 %	89 710
7/a.	Transport	-0.00%	3.75 %	47 817

Statistics

Updated: 2022-11-14 **Evolution over time Current State**

Evolution over time

Total Number of Datasets Datasets by data theme Datasets per origin Datasets per origin and catalogue









Country Insights

This section provides insights into the status of open data in European countries. The section includes information about countries' national open data portal, the level of open data maturity, use cases, catalogues, events, reports, news, webinars, and interviews about open data. Read more...



Open Data Impact

The 2020 report on the Economic Impact of Open Data researches the value created by open data in Europe. It is the second study by the European Data Portal, following the 2015 report. The study forecasts the open data market size and the employment growth for 2025. Impact is exemplified by efficiency gains and cost savings due to open data. Read more...



Open Data Maturity

A series of indicators have been selected to measure open data maturity across Europe. These indicators cover the level of development of national policies promoting open data, an assessment of the features made available on national data portals as well as the expected impact of open data. Read more...



Studies

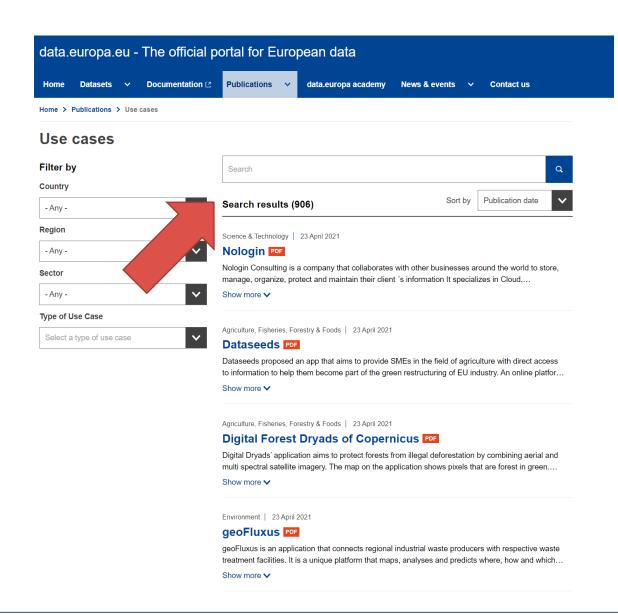
A collection of papers that discuss various relevant topics around open data, such as barriers in working with open data, high-value datasets, and business-togovernment data sharing. Read more...



Use Cases

This part of the portal contains use cases from both publishers and re-users of open data. The use cases provided below are gathered via the 'Tell us your story' form on the portal or shared with us by other contacts. This section of the portal will grow over time. Read more...













Log in English

data.europa.eu - The official portal for European data

Home Datasets ∨ Documentation ☑ Publications ∨

data.europa academy News & events 🗸

Contact us

Home > Publications > Open Data Maturity

Open Data Maturity

The open data maturity study serves as a benchmark to gain insights into the development achieved in the field of open data in Europe. It assesses the level of maturity against four dimensions; policy, portal, impact, and quality. The study clusters countries into four different groups: trend-setters, fast-trackers, followers, and beginners, from the most mature to the least. Recommendations tailored to the level of maturity and characteristics of each of these groups are provided. Read the latest report here

2021	2020	2019	2018	2017
2016	2015 PDE			

Dimension	Metrics		
	Policy framework		
Open Data Policy	Governance of open data		
	Open data implementation		
	Strategic awareness		
	Political impact		
Open Data Impact	Social impact		
	Environmental impact		
	Economic impact		
	Portal features		
Owen Dete Destel	Portal usage		
Open Data Portal	Data provision		
	Portal sustainability		
	Currency		
Out of Date Out lite.	Monitoring and measures		
Open Data Quality	DCAT-AP compliance		
	Deployment quality and linked data		

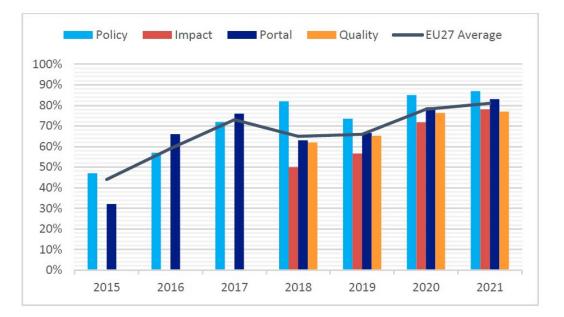


Figure 1: The Open Data Maturity scores of the EU27 (used to be EU28 until 2019)







data

europa

academy

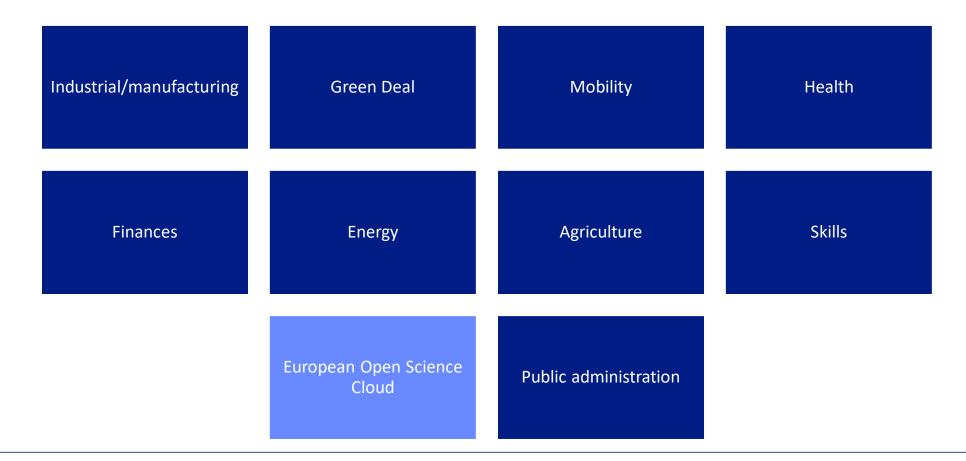


data.europa.eu and data spaces





Common European data spaces (initial list)

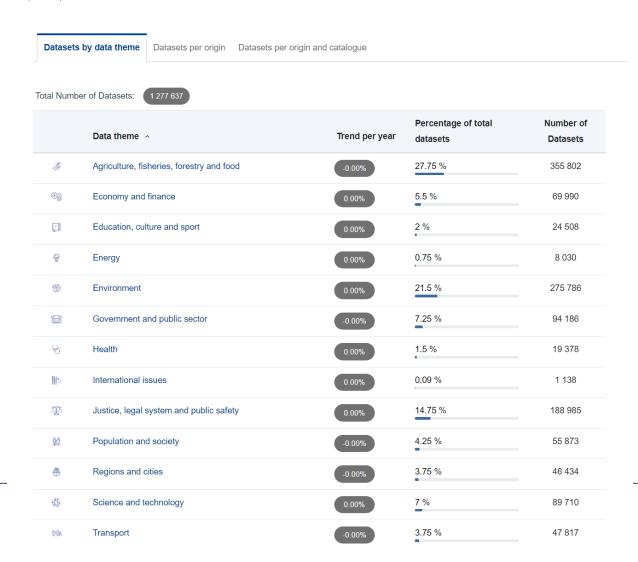






Premise of the research (1)

 data.europa.eu could be a core supplier of public sector data for European common data spaces







Premise of the research (2)

- Ongoing initiatives and funded projects will generate lists of datasets relevant for each data space and governance schemes to connect existing data ecosystems.
- Public administrations and other open-data holders/publishers must be part of this conversation.





Goals of the research

- Open government data (OGD) should not be an afterthought to the European single market for data
- Two reports
 - This report (completed in Q1 2022)
 - Explore challenges and opportunities for data.europa.eu in the context of data spaces at regional, national, and European levels
 - Landscaping and analysis of ongoing works towards standards, consolidated support, and implementations
 - Next report after (some) spaces become available
 - data.europa.eu as intermediary in existing spaces
 - Technical roadmap, synergies, collaborations





Methodology

Desk research **Interviews**

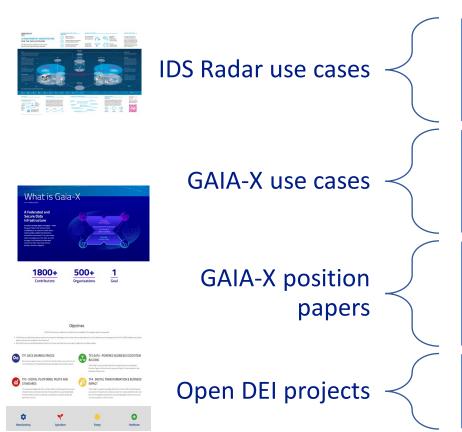
- Identify data spaces
- Identify use cases and open data in these use cases
- Map out PS actors involved
- Select stakeholders to interview

 Add context and depth to desk research findings





Desk research: sources



- 57 use cases: smart cities (4), manufacturing (7), energy (4), mobility (7), automotive (4), supply chains (8), cross-sector/other (24)
- 78 use cases: agriculture (4), energy (9), finance (3), geoinformation (5), health (22), industry 4.0 (14), mobility (5), public sector (10), smart cities (1), smart living (5).
- Agriculture, energy, finance, geoinformation, health, industry
 4.0, mobility, public sector, smart cities/smart regions, smart living
- 34 projects: manufacturing (7), agriculture (6), energy (8), health (13).





Desk research: methods

- Check if they mention or use open data from any source or portal
- Check if they engage actively with key open-data stakeholders e.g.
 PS publishers, data intermediaries.

- Discarded all spaces (and related use cases) that referred to:
 - Tech for data spaces rather than actual data spaces;
 - Enterprise data spaces benefiting individual organisations.





Interviews

- 12 participants from 8 countries (19 people contacted)
- 30 minutes semi-structured interviews

Id	Domain	Country	Type of stakeholder	Gender
P1	Mobility	Spain	Data owner Data user	Male
P2	Health	France	Software provider	Male
Р3	Cross-domain	Germany	Software provider	Male
P4	Public sector	Belgium	Vocabulary provider	Female
P5	Industry 4.0	Spain	Data provider	Prefer not to say
P6	Geoinformation Public sector	Italy	Service provider	Male
P7	Public sector	Italy	Broker service	Female
P8	Agriculture	Poland	Vocabulary provider Data provider	Female
P9	Cross domain	Spain	Identity provider	Male
P10	Cross domain	France	Software provider	Male
P11	Mobility	France	Data provider	Female
P12	Supply Chain and Logistics	The Netherlands	Vocabulary provider Male	

ANNEX 1. Interview protocol

This annex provides the interview protocol that has been used for the semi-structured interviews that have been done in the context of this work. As in any semi-structured interview, the questions that are presented here are mainly acting as a guideline for the conversation with the interview.

1. Are you involved in the development of any emergent or well-established data space?

If yes, in which domain or group of domains?

2. Are you involved in the setup of any of the European common data spaces?

If yes, in which one? Are you submitting or have submitted a request for funding under any of the Digital Europe or Horizon Europe programme calls?

3. Do you or the organisation that you represent belong to any of the organisations that are active in the development of frameworks and architectures for data spaces (IDSA, GAIA-X, etc.)?

4. For each of the data spaces that you are involved in, can you tell me...? (no need for specific names, the answers can be similar to "a relevant company in the automotive sector")

4.1 The private organisations that are involved on it

4.2 The public organisations that are involved on it

4.3 Any other type of organisation or individual that is involved on it

5. Do you already have an initial catalogue of datasets identified for your data space?

5.1 How many of these datasets are from public administrations?

5.1.1 If there are public datasets, are the public administrations that act as data providers or data intermediaries aware of the fact that these datasets will be used in the data space?

5.1.2 Have you thought of using data.europa.eu or any other open data portal as a source for data?

5.1.3 Will the data space contribute any data to data.europa.eu or any other open data portal?

5.2 How many of these datasets are from private institutions?

5.3 Are there any other relevant set of datasets that will be used?

6. Have you decided on the metadata schemas that will be used to describe (and federate) datasets? Something like DCAT or the IDS Information model?

7. Do you know of other similar data spaces in the domain of your data space that we should look at?

8. Would you have any recommendation for data.europa.eu on how to make sure that their datasets are used in the data space?





151 initiatives, 21 use open data, 2 mention data.europa.eu, 19 involve PS stakeholders

Domain	Initiatives where open data is used	Initiatives where data.europa.eu is mentioned	Initiatives with public stakeholders involved	Total number of initiatives
Agriculture/Agrifood	2	1	0	9
Energy	2	0	3	21
Finance	1	0	1	3
Geoinformation	5	1	0	5
Health	0	0	7	37
Industry 4.0/ Manufacturing	0	0	0	27
Automotive	0	0	0	4
Supply Chain and Logistics	0	0	0	8
Mobility	5	0	4	10
Public Sector	1	0	1	9
Smart Living / Smart Cities	5	0	3	10
Other	0	0	0	8
Totals	21	2	19	151





Overlapping priorities in data space development

European common data spaces (source: Digital Europe work programme)	Domains for data spaces identified by GAIA-X, IDSA and Open DEI
Green Deal	Geoinformation
Smart communities	Smart living and smart cities
Mobility	Mobility
Manufacturing/Industrial	Industry 4.0/Manufacturing/Automotive
Agriculture	Agriculture / Agri-food
Cultural heritage	-
Health (focus on cancer and genomics)	Health
Media	
Finance	Finance
Skills	-
Language	-
Public procurement	Public sector (procurement not explicitly considered)
Security and law enforcement	Public sector (procurement not explicitly considered)
Tourism	-
Energy	Energy
-	Supply chains and logistics





Additional general findings

- Bias on French and German organisations in use cases, data sources, stakeholders, though developments under way to change that e.g. GAIA-X AISBL.
- Three sectors (geo, mobility, smart cities) with clear ties to open data sources, though no portals mentioned, or PS stakeholders mentioned or involved.
- No catalogues of datasets in either source consulted.





Sector-specific findings (1)

- Agriculture: Earth Observation data, INSPIRE, no PS involvement.
- Energy: municipal open data for business models in energy, cities, national agencies.
- Finance: financial big data cluster with open data component, several PS administrations as stakeholders.
- Geoinformation: various open data sets but no specific providers or PS stakeholders involved.
- Health: open data only for Covid-19 dashboard/hub, public healthcare providers.





Sector-specific findings (2)

- Manufacturing, automotive, supply chains, logistics: no open data references or PS stakeholders involved.
- Mobility: most PS stakeholders involved from all sectors analysed, no specific data portals, catalogues etc.
- Public sector: no PS stakeholder involvement
- Smart cities, smart living: some data providers in the public sector identified, but not involved.





Interview findings (1)

 PS organisations are underrepresented in ongoing data sharing initiatives.

"The role of **open data portals as data providers is taken for granted** in many of the data spaces (and use cases) that are being developed. However, the institutions in charge of open data portals should be more actively involved in the development of data spaces in order to have more representativeness in those developments". (P6, provides services to public administrations in Italy)





Interview findings (2)

 No catalogues of datasets to be covered in the spaces.

"Open data portals like data.europa.eu have demonstrated the possibility of federating metadata and providing links to datasets. This type of technology is clearly useful for many data spaces that do not have additional requirements for the collection of the catalogue of datasets, so institutions like data.europa.eu should make an effort to position their open source technology in as many data spaces as possible". (P3, develops architectures and technology for spaces, Germany)





Interview findings (3)

 Underspecified metadata schemes, identifier schemes, vocabularies etc.

"Public administrations (e.g., those behind data.europa.eu) may act in some occasions as **neutral organisations that can provide support for some of the roles identified in the main architectures**. For instance, they may have a relevant role as **vocabulary providers**, including the participation in the development of **data standards and models, or even as identity providers** or as part of the data governance managers." (P12, develops vocabularies for supply chain and logistics, The Netherlands)





Conclusions and next steps





Conclusions

- Data spaces is an emerging field, with activities spanning across a range of stakeholders. Still some biases and limitations could be identified, both geopolitical and in terms of stakeholders considered or directly involved.
- Open government data sources, portals, stakeholders are not as prominently represented as they could be, not even in sectors where they play a key role: mobility, smart cities, smart communities etc.
- Current initiatives do not consider some spaces that are relevant to the public sector e.g. procurement, skills, cultural heritage, green deal.





Opportunities will be missed

- Open data is a core type of data source, but unless PS stakeholders and open data intermediaries are involved more directly in IDSA, GAIA-X, Open DEI, there will be frictions and open data will not have the impact it deserves.
- Experience of open data holders with data publishing, metadata management, data quality, data discovery, federation etc. is not transferred to data spaces. Key European expertise and technologies have less impact.
- Interoperability requires neutral entities to define identifies, exchange formats, vocabularies. PS data intermediaries like data.europa.eu are a natural choice to take on such roles.





Next steps

- Once some data spaces become available, undertake in-depth case study analysis focused on data.europa.eu as a data intermediary.
- Derive technical and non-technical roadmap to allow data.europa.eu to be used effectively in data spaces.





Thank you

info@data.europa.eu





