Creating public sector value through the use of open data

Insights and recommendations from the data.europa.eu campaign

Summary paper 2023





This study has been prepared as part of data.europa.eu. Data.europa.eu is an initiative of the European Commission. The Publications Office of the European Union is responsible for contract management of data.europa.eu.

For more information about this paper, please contact:

European Commission Directorate-General for Communications Networks, Content and Technology Unit G.1 Data Policy and Innovation Email: <u>CNECT-G1@ec.europa.eu</u>

data.europa.eu Email: <u>info@data.europa.eu</u>

Written by: David Osimo Anna Pizzamiglio

Last update: 8 August 2023 https://data.europa.eu/

Disclaimer

By the European Commission, Directorate-General of Communications Networks, Content and Technology. The information and views set out in this publication are those of the author(s) 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 which may be made of the information contained therein.

Luxembourg: Publications Office of the European Union, 2023 © European Union, 2023



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

ISBN: 978-92-78-43685-8 doi: 10.2830/565958 OA-04-23-813-EN-N

Creating public sector value through the use of open data

This summary paper, *Creating public sector value through the use of open data*, is the last of two papers aimed at fostering the engagement of public sector as reusers of open data. It sheds light on the results from the 1-year campaign featuring a first scoping <u>paper on measuring data demand within the public sector</u>, several <u>blog posts</u> published on data.europa.eu, a <u>data.europa.eu academy webinar</u> on data demand and a focus group on how data is reused by the public sector.

The summary paper incorporates the insights from the campaign, draws conclusions and provides policy recommendations on how data.europa.eu can enable and stimulate data reuse by public institutions across Europe.

~				
Со	n	tρ	n	t
0	•••	L C	•••	•

Executive summary	5
1. Introduction	6
1.1. Putting the spotlight on public institutions as open data reusers	6
1.2. Aim and structure	7
2. Untapped potential: how public institutions can benefit from reusing government open data	8
2.1. How important is public sector data reuse? Some data points	8
2.2. Good practices of data reuse by public institutions in Europe	9
2.2.1. National example: France	10
2.2.2. National example: Estonia	13
2.2.3. Regional example: Flanders, Belgium	15
3. Barriers to data-driven value creation in the public sector	17
3.1. Technical barriers	17
3.2. Institutional barriers	19
4. Empowering public institutions to reuse open data	21
4.1. Raising awareness	22
4.2. Policy and regulatory frameworks	22
4.3. Monitoring	22
4.4. Data stewardship	23
5. Conclusions and insights for data.europa.eu	23
6. Bibliography	24
7. Annex: Methodology	24
7.1. Mini-survey structure	25
7.1.1. Mini-survey n1	25
7.1.2. Mini-survey n2	25
7.2. Webinar: data demand and reuse in the public sector	26
7.3. Focus group: how open data is used by the public sector	26
7.4. Expert interviews structure	27

Executive summary

This is the final paper on public institutions as reusers of open data. It follows the 2021 discussion paper, which focused on how to measure data demand within the public sector. While the first paper considered a primary stage in the open data value chain, this summary paper looks at the final stage, i.e., the creation of value by the public sector through the use of open data, and is structured around the following question:

How to foster open data value creation in the public sector?

This question is considered from a practice-oriented angle, supported by the data.europa.eu community engagement activities that have taken place through a campaign since the publication of the discussion paper. This summary paper provides an overview of the different stakeholder activities undertaken, ranging from surveys to a focus group, and presents the key insights from this campaign regarding data reuse practices, barriers to data reuse in the public sector and suggestions to overcome these barriers.

The following recommendations are made to help data.europa.eu support public administrations to boost open data value creation.

- When it comes to <u>raising awareness and communication</u>, any action should also contain examples of data reuse by the public sector. Gathering and communicating such examples and use cases greatly helps in understanding the importance of the role of the public sector as a data reuser.
- When it comes to <u>policy and regulation</u>, it would be beneficial to align the 'better regulation' activities and roadmaps of the European Commission with the open data publication activities, in order to better explore the internal data needs. Furthermore, it would be helpful to facilitate a similar alignment and data needs analysis for all European public administrations. For example, this could be done by providing examples, best practices and methodologies on how to map data needs for policy and regulatory purposes.
- Existing <u>monitoring</u> activities, such as surveys, should be revised to ensure that data reuse by the public sector is included. It would be useful to create a panel of users, based on the existing wide community, that could be used for further surveys.
- The role of <u>data stewards</u> remains central to favouring reuse. Therefore, examples, best practices and methodologies on the role of data stewards should be included in the support activities not specifically for public sector reusers, but in general.

1. Introduction

1.1. Putting the spotlight on public institutions as open data reusers

As outlined in the discussion paper 'Measuring data demand in the public sector', most data reuse studies and initiatives have been directed towards the private sector reusing government data ⁽¹⁾. This is because the potential for economic value creation through open data policy became evident and the private sector was considered to be the primary beneficiary. The underlying assumption is that the private sector leverages government data to create services that generate higher returns in terms of public value and revenue. On the other hand, limited attention was given to exploring the value of open government data for the public sector.

The discussion paper also underlined that data reuse represents an invaluable resource for the public sector. By leveraging data, not only can services and policies become more efficient, but also more sustainable, inclusive and trustworthy, thereby directly benefiting citizens and businesses. European governments are thus increasingly exploring how to foster a data-driven public sector that recognises data, to varying degrees of openness, as an asset integral to policymaking, service delivery, organisational management and public innovation ⁽²⁾.

However, to reap such benefits, it is not enough to increase efforts at the start of the data value chain by measuring public institutions' data demand and publishing data accordingly. Besides understanding the data needs of public institutions, it is necessary to understand what hinders them from using the data at the end of the value chain, thereby creating public value. There are still barriers that prevent public institutions from seizing the full potential of data-driven policies, including technical, cultural and regulatory factors. To increase the understanding of these issues, explore ways to address them and engage the data.europa.eu community on this topic, several activities were undertaken as a campaign to raise awareness of the value of open data within data.europa.eu. Extracting insights from this campaign is crucial to enable data.europa.eu to play its role in facilitating open data reuse by public institutions at all government levels across Europe.

The data.europa.eu campaign (Figure 1) consisted of a discussion paper, a webinar, two mini-surveys, a blog series and a focus group, all of which provided valuable insights into the challenges and opportunities of data-driven value creation in the public sector. The mini-survey and focus group involved public institutions across Europe, while the webinar and blog series targeted a wider audience of data professionals and enthusiasts (more detailed explanation available at the <u>Annex</u> of this study). Reflecting on these activities and the collected insights regarding user engagement and experiences, will shape how to foster the uptake of data reuse going forward for data.europa.eu.



Figure 1: The data.europa.eu campaign

^{(&}lt;sup>1</sup>) van Ooijen, C., da Silva Carvalho, N., Iordache, A. et al., 'Measuring data demand within the public sector: Discussion paper 2021', Publications Office of the European Union, Luxembourg (<u>https://data.europa.eu/doi/10.2830/54555</u>).

^{(&}lt;sup>2</sup>) van Ooijen, C., Ubaldi, B. and Welby, B. (2019), 'A data-driven public sector: Enabling the strategic use of data for productive, inclusive and trustworthy governance', OECD Working Papers on Public Governance, No. 33, OECD Publishing, Paris (<u>https://doi.org/10.1787/09ab162c-en</u>).

1.2. Aim and structure

This summary paper builds on the work done within the scope of the data.europa.eu project to facilitate the exchange of examples of data reuse value creation by public institutions at the EU, national and local levels. It analyses and discusses how countries involve potential reusers and how they promote reuse within governments at all levels.

The summary paper captures the insights from the data.europa.eu campaign, thereby providing a basis for data.europa.eu to:

- 1. improve the tailoring of existing tools, resources and support mechanisms to the needs of:
 - i) public institutions as data providers wishing to foster the reuse of their data by other public institutions,
 - ii) public institutions as potential reusers of open data;
- 2. develop new tools, resources and support mechanisms to better respond to the needs of the aforementioned target groups.

To guide the research activities supporting the drafting of the summary paper, the following research question and sub-questions have been formulated:

How can data-driven value creation in the public sector be fostered?

This main question is answered by addressing both the supply and demand sides, as elaborated in the following subquestions.

- To what extent do public institutions use open data from other public institutions, and how is this data being utilised?
- What are the current barriers to leveraging open data for data-driven value creation in the public sector?
- What are the enabling factors that can stimulate open data reuse by public institutions across Europe?
- What lessons can be learned from successful public sector data reuse initiatives regarding the needs of public institutions as data reusers?
- How can data.europa.eu effectively support public institutions, both as data providers and reusers in realising the potential of data-driven value creation?

The summary paper is organised into four main parts, corresponding to and aiming to address each research question.

In the second chapter, what we know about open data reuse by the public sector will be outlined. This includes an illustration of data reuse practices by the public sector, explaining how data reuse takes place. Specific illustrative cases gathered in the course of the study will be presented.

In the third chapter, an analysis of the several challenges and barriers to data-driven value creation in the public sector revealed through the data.europa.eu campaign is presented. These challenges include lack of awareness and understanding of the potential of data, insufficient data quality, interoperability and accessibility issues and limited human and financial resources.

In the fourth chapter, the report outlines existing practices to favour data reuse by the public sector. Among the factors explored are demand-driven data publication, data governance frameworks, data sharing and collaboration platforms and capacity building and training initiatives.

Finally, the report presents concluding recommendations for the data.europa.eu portal and stakeholders in general. The recommendations will serve as a roadmap for data.europa.eu and other stakeholders to enable and stimulate open data reuse by public institutions across Europe and to realise the full potential of data-driven value creation.

2. Untapped potential: how public institutions can benefit from reusing government open data

This section of the paper will illustrate the limited quantitative evidence available concerning the impact and importance of data reuse by the public sector, as well as virtuous cases of data reuse by European public institutions gathered during these activities.

2.1. How important is public sector data reuse? Some data points

The examples show the wide variety of use cases of open data reuse by the public sector. But open data reuse is not yet a well-established practice across the public sector in the European Union. What do we know about the overall importance of the public sector as a reuser? How widespread is open data reuse by the public sector?

Answering this question is surprisingly challenging. There are no dedicated studies on this topic and empirical data is scarce. Throughout the wide-ranging review conducted as part of this campaign, one of the few hard data points that were found is part of the annual survey of open data portals in Spain, which is carried out every 2 years. The results should be taken with a grain of salt, as they are based on the perception of the person in charge of the open data portal rather than on actual data. Interestingly, by far the top reuser of open data is the actual organisation running the open data portal – which is a frequent reuser of open data in 63.9 % of cases. However, other public institutions were frequent reusers only in 13.9 % of cases in 2021 ⁽³⁾.

The mini-surveys carried out as part of the data.europa.eu campaign, while not being statistically representative, offer new data points that provide a useful snapshot of the perception of the community. The purpose of the surveys was to better understand the demand for open data among public administrations in Europe and the challenges they face in adopting and implementing open data policies. We surveyed a diverse range of public administrations across Europe, including national, regional and local governments, along with public institutions and agencies, and we collected a total of 108 answers to the first survey and 548 answers to the second one.

When respondents were asked about the importance of demand for open data by the public sector, a staggering 92.52 % indicated it was high or very high. Furthermore, when asked about whether their administrations use open data from other government entities, 59.49 % answered positively. These two data points alone confirm the importance of open data reuse by the public sector. Moreover, all experts involved in this campaign confirmed this assumption, through interviews, webinars and focus groups. As one expert stated at the very beginning of the 'Data demand and reuse in the public sector' webinar, hosted in March 2022 by the data.europa.eu academy, 'There is this fantasy that open data is made primarily for start-ups, but at least in France the first users of open data are public bodies.'

^{(&}lt;sup>3</sup>) Abella, A., Ortiz de Urbina Criado, O., De Pablos Heredero, C. and García Luna, D. (2022), 'Reusing open data in Spain III. October 2021', ESIC EDITORIAL, Madrid



Figure 2: Visualisation of the mini-survey n2 results

The March 2022 webinar also saw Frederika Welle Donker, researcher at the Knowledge Centre Open Data of Delft University of Technology, giving an overview of the research they carried out in 2019 for the Dutch Ministry of the Interior and Kingdom Relations into open data reusers and illustrating developments seen since then. The study was commissioned because the ministry had indications that open data was being reused on a large scale, but had no idea by whom and for which goals, and wanted to gain some insights into that.

The research consisted of both a quantitative analysis, i.e., analysing IP log file numbers from Dutch open data portals, and a qualitative analysis, i.e., posting an online questionnaire via social media and organising research with Delft University researchers. After the quantitative analysis, they selected reusers for interviews to validate the findings. The analysis of log files indicated that **the majority of users belonged to the public sector.** Further, the study revealed that public-sector users demonstrated a greater awareness of who the data holders are and tended to establish direct communication with them. In contrast, private-sector users and citizens relied more on portals, platforms, search engines like Google or their personal networks to locate the desired data.

Additionally, Ms Donker highlighted that following the research work – and possibly in the aftermath of the COVID-19 crisis – there **has been a shift from seeing open data as a goal to using it as a means for digital transformation**. Additionally, there has been a transition from data portals to more interactive data platforms, possibly in anticipation of upcoming EU legislation or as part of ongoing developments. However, it is important to note that the research was conducted in 2019, i.e., before the publication of the European data strategy.

The lack of policy reports and academic publications on European institutions' open data activities and efforts to measure public institutions' data demand does not mean that European data providers do not engage with public sector users or other communities. It simply suggests that such activities are under-documented, at least in the public sphere.

2.2. Good practices of data reuse by public institutions in Europe

Today's society is based on the mass processing of data and public institutions play a significant role as data providers. Many of the data can be reused by public institutions to create value, not just from an economic perspective but also in the social and environmental spheres. The Organisation for Economic Co-operation and

Development (OECD) has identified three main areas in which data-reuse can enhance the effectiveness of policymakers and public officials across policy areas and government levels ⁽⁴⁾.

The first is **anticipatory governance**, which refers to systematic efforts by governments to consider the future and inform policy decisions today. Open data can be used to predict trends and patterns in order to mitigate emerging risks and respond to developing crises.

A second area of opportunity is the **design and delivery** of policy and services, where data provides important ways of understanding problems, engaging the public and providing access to insights for improving public services that meet user needs, while creating the conditions for robust, evidence-based policy making.

Thirdly, data can be leveraged for **performance management** in terms of greater public sector productivity, more efficient use of resources and better evaluation of policies and impact.

Anticipatory governance

- Predict trends and patterns
- Mitigate emerging risks
- Respond to developing crises

Design and delivery

- Understanding problems
- Engaging with the public
- Meeting citizens' needs
- Evidence-based policy

Performance management

- Public sector productivity
- More efficient use of resources
- Evaluation of policies and impact

Figure 3: Areas in which data reuse can enhance effectiveness

In the context of the 'Data demand and reuse in the public sector' webinar, hosted in March 2022 by the data.europa.eu academy, and during the focus group on 'How open data is used by the public sector', hosted in March 2023, participants had the opportunity to share what their governments are doing and how they are exploiting open data. Several examples of good data reuse practices from European countries were presented.

2.2.1. National example: France

Antonin Garrone, product owner at Etalab, the French task force for data policy including open data, discussed fostering data reuse in the public sector and highlighted three main use cases to illustrate data reuse in France.

(1) **Public services.** Mr Garrone mentioned digital public services that produce, exploit and distribute open data as part of their value proposition. For example, <u>acceslibre</u> (free access) gathers information and maps the accessibility of establishments open to the public. It aims to identify the accessibility of places receiving the public and to share this information as widely as possible. Additionally, citizens and shop owners can add information essential to the movement of people with disabilities. There is also the <u>Dataqir</u> platform, which develops simulators that evaluate the environmental impact of behaviours such as food consumption and mobility. These datasets are available for reinterpretation and incorporation into other systems. *Datagir* and *acceslibre* are clear examples of how open data reuse by public institutions can help in the **design and delivery of public services.**

⁽⁴⁾ See footnote 2.





(2) The <u>Base Adresse Nationale</u> (BAN, National Address Database) associates each address listed on French territory with its geographic coordinates. It does not contain any nominative data. It was built from address databases of La Poste, the National Geographic and Forest Information Institute (IGN) and the Public Finances General Directorate. It was then enriched by data produced by governmental agencies and is further developed by citizen contributions. On **adresse.data.gouv.fr**, municipalities, enterprises and citizens can contribute and freely use this database on the principles of collaboration (share, improve and reuse a common good). The following three levels of services are offered to individuals, enterprises and local governments:

- free download of address files;
- the first French geocoder, comparable to web tools provided by large internet players;
- a *Guichet Adresse Mairie* (local address counter) to support municipalities in creating, identifying and numbering the road network.

The BAN is a crucial dataset in France and is used in both private and public applications. It enables features like autocomplete when entering addresses and facilitates the installation of broadband networks. Moreover, it assists emergency services in accurately locating firefighter responses. The BAN is an essential resource for various sectors that rely on precise address information. Considering its various uses, the tool allows for both **anticipatory governance**

and **design and delivery** (e.g., it shows where high-speed internet connections need to be improved or made available).



Figure 5: Functions available at adresse.data.gouv.fr

(3) The <u>Annuaire des Entreprises</u> (enterprise directory) is a website that allows you to find all the public data held by the French administration on a company, an association or an administration. While the site mainly acts as a centralised hub for enterprise-related data, it also serves as a cornerstone for multiple public services in France and supports various applications. For instance, the directory has developed its own application programming interface (API) for its search engine and makes it available free of charge.



Figure 6: Annuaire des Entreprises' API search engine

2.2.2. National example: Estonia

Estonia is another clear example of how public authorities have begun the process of shifting their mindset from being solely data publishers and holders to becoming data users and reusers since the COVID-19 pandemic. During the focus group discussions, a representative from the Ministry of Economics and Communications shared concrete examples of how open data is used in Estonia.

Firstly, public authorities and citizens can use an interactive dashboard called <u>Tõetamm</u> (Tree of Truth) to track the progress of government action in specific policy areas. The dashboard showcases progress in different regions on indicators that are vital for the country's development, including national security and defence, education, welfare and the rule of law. The Tree of Truth represents the 135 most important indicators, as defined by the Estonian main government office, and which originated from three different government development plans: (1) Estonian development plan goals for 2020; (2) Estonian long term development plan for 2035; and (3) sustainable development goals.



Figure 7: Estonia's Tree of Truth

The Tree of Truth is a customised platform that was designed to meet the needs of Statistics Estonia. Every leaf represents an indicator and the tree as a whole shows the health of Estonian society. If you click on a leaf, there emerges more detailed information for each indicator, explaining the value and showing values from previous periods for comparison. The aim of the dashboard is to show the 135 indicators in a single comprehensive view. The Tree of Truth is an example of how insights based on government data can be crucial in understanding problems, engaging the public and providing access to insights for improving public services that meet user needs, while creating the conditions for robust, evidence-based policymaking.

Secondly, decision-makers and citizens in Estonia have access to various <u>dashboards</u> that show indicators on different fields. They can compare these fields and make decisions based on the data presented.

At least 15-year-o persons by educa At least 15-year-old per and sex I Whole country	sons by educational level	Deaths Deaths by sex I Whole of	country, 2015–2022	Demographic labour pressure index Demographic labour pressure ind country, 2017–2022	dex I Whole
240,000		20,000		1	
180,000		15,000		0.75	
120,000		10,000		0.5	
60,000		5,000		0.25	
0		0		0	
2018 2019	2020 2021 2022	2015 2016 2017 2	018 2019 2020 2021 2022	2017 2018 2019 2020	0 2021 2022
	POPULATION		POPULATION		POPULATION
Dependency ratio	₽ ↓ ×	Emigration	∓	Immigration	∓
Dependency ratio I Who	ble country, 2017–2022	Emigration by type of m country, Males and fema		Immigration by type of migration country, Males and females, 201	
60		14,000		60,000	
45		10,500		45,000	_
30		7,000		30,000	
15		3,500		15,000	
		0		0	
0					
0 2017 2018 2019	2020 2021 2022	2015 2017	2019 2022	2015 2017 2019	2022



Thirdly, Estonia leverages open data to create artificial intelligence and data-based solutions **by collaborating with other state authorities** ⁽⁵⁾. The Ministry of Economic Affairs and Communications uses the <u>Kaggle</u> platform to develop data-based solutions that address existing issues. This is an apt example of how open data can help with anticipating trends and patterns in order to mitigate emerging risks and respond to developing crises, i.e., an instance of how data can help with **anticipatory governance**.

2.2.3. Regional example: Flanders, Belgium

Another example of data reuse that facilitates **anticipatory governance** is that of Flanders in Belgium. The website <u>provincies.incijfers.be</u> is a collaboration between the data and analysis departments of the five Flemish provinces.

In each province, these departments provide **support in policy planning** to local administrations and organisations through figures and reports on numerous themes and policy areas, mainly through the dashboard and database of provinces.incijfers.be. Local authorities and organisations can contact the departments for questions about figures, for (environmental) analyses and for interpretation of figures. Through provinces.incijfers.be, the departments are constantly expanding the structural supply of figures and reports. This is done in cooperation with and at the request of other provincial departments, external partners (such as the Knowledge Centre for Flemish Cities) and local authorities. In doing so, they use figures from official data sources, which are made accessible in a user-friendly way at different area levels, from statistical sectors to regions.

^{(&}lt;sup>5</sup>) <u>https://e-estonia.com/come-and-help-to-solve-real-world-ai-projects-on-kaggle/</u>



Figure 9: Examples of dashboards available on provinces.incijfers.be

In conclusion, the findings show that while public administrations recognise the potential benefits of open data, little action is taken to explicitly target and measure the reuse of data by other governmental entities. Nevertheless, authorities are already using open data to design policies and services for their citizens. The examples shared during the 'Data demand and reuse in the public sector' webinar and the 'How open data is used by the public sector' focus group demonstrate how open data is being leveraged in the public sector to enhance services, improve accessibility and enable data-driven decision-making. By sharing and reusing data, public bodies can drive innovation, foster collaboration and meet the evolving needs of citizens and organisations.

3. Barriers to data-driven value creation in the public sector

To unlock the potential of open data reuse in the public sector, it is crucial to address challenges such as technical difficulties, administrative delays and legal and financial restrictions that can hinder data accessibility and reusability. Failure to address these challenges may impede the achievement of the benefits of national efforts aimed at promoting reuse of open government data. To overcome these obstacles and develop sustainable business models for open data reuse that can generate the desired benefits, it is essential to address various challenges related to policy, technology, financing, organisation, culture and legal frameworks.

The analysis of the data.europa.eu campaign revealed several issues when it comes to fostering open data reuse within the public sector. These can be divided in two broad categories: technical barriers and institutional barriers.



Figure 10: Mini-survey n1 – barriers to data-driven value creation

3.1. Technical barriers

Value creation through the reuse of open data by public institutions faces different technical barriers. The evidence gathered during the data.europa.eu campaign suggests that the most common ones are:

- data silos, and the related lack of data harmonisation and interoperability;
- poor data quality (and management);
- licensing issues.

Data silos have been a longstanding practice for many organisations and institutions. They were the 'old' way of storing data, where each department or system within an organisation had its own isolated repository of data. However, the shift towards open data sharing has highlighted the limitations of this approach.

First, data silos create obstacles to data sharing and collaboration by fragmenting data, leading to duplication and inconsistencies. These isolated repositories make it difficult to integrate and combine data from different sources, hindering a comprehensive understanding of the data. Moreover, limited data accessibility within silos, due to access controls and proprietary formats, restricts collaboration and informed decision-making. Second, managing data silos incurs costs and inefficiencies, as each silo requires its own infrastructure and management processes. Additionally, the lack of interoperability between silos leads to inefficient data integration and analysis. This was confirmed during the focus group discussion. A first-hand example of this was relayed by a data expert from Malta, where API sharing practices were introduced almost a decade ago, which is a positive sign of data-driven policymaking. However, with the emergence of open-source systems, it has become difficult to adjust the existing system to reflect these technological advancements. The migration from one system to another has proven to be particularly challenging. National experts who participated in the focus group and expert interviews reported that there is still some resistance to changing legacy applications. This resistance to change can be a hindrance to the adoption of new technology and affects the ability of public institutions to fully achieve the potential of data.

Data quality is another crucial challenge that hampers the effective use and reuse of open data. Poor-quality data is inaccurate data. Examples include missing contact fields (phone numbers, emails, physical addresses, etc.), outdated information (old job titles, changes caused by mergers and acquisitions, etc.), data entered in the wrong field, duplicate entries, misspellings and spelling variations and non-normalised data. Poor data quality can result from various factors, such as errors during data entry, inadequate data validation processes, outdated or obsolete information, lack of data governance or problems with data integration from multiple sources. Without continuous, real-time maintenance, data can stagnate and decay, making accurate database maintenance a difficult task. Databases must be cleaned, maintained and appended regularly. Ensuring high-quality data is essential for maximising the value and potential of open data initiatives. Efforts must be made to improve data quality through standardised data collection processes, quality control measures and regular data maintenance and updates. By addressing data quality challenges, public institutions can enhance the reliability and usefulness of open data, fostering its effective reuse and driving positive outcomes.

Another key issue raised by almost all of the national data experts present in the focus group is **licensing**. One of the core principles of open data is the ability to reuse and redistribute the data freely. However, licensing requirements can undermine this principle, limiting the reusability and redistribution of open data. These limitations can impede the potential for collaboration, innovation and the reuse of open data. In the context of the focus group discussions, participants highlighted that defining what truly constitutes open data can be problematic. Many administrations publish APIs, but the issue lies in the licensing requirements, which often necessitate subscription or access fees. This situation indicates that open data may not be as open as it should be. Additionally, it was highlighted that in order to achieve genuine openness, there is a need to streamline and align the data, ensuring its proper publication. The restriction on the redistribution and reuse of data due to licensing requirements often undermines the idea of open data itself.

Moreover, licensing can cause compatibility issues. Different datasets may be subject to various licensing terms and conditions, which can create compatibility issues when trying to integrate or combine data from multiple sources. Inconsistent licensing requirements make it difficult to create comprehensive datasets or to develop tools and applications that utilise data from different providers. This fragmentation hampers interoperability and reduces the usability of open data.

Another crucial barrier to open data reuse arising from licensing is the administrative and bureaucratic burden required to set the right licensing system. Indeed, licensing agreements impose administrative burdens, such as seeking permissions, negotiating terms or seeking legal advice. These requirements create barriers, particularly for smaller organisations or individuals with limited resources, as the time, effort, and cost involved in navigating licensing processes can discourage data sharing and hinder the adoption of open data practices. An expert from Opendata.fi, the open data portal of the Finnish Digital Agency, illustrated this challenge during the focus group by citing the example of the Finnish service catalogue, where local authorities repurpose open data. While the register built upon this data showcases the value of using open data, the establishment of licensing mechanisms to govern data usage becomes a significant consideration. He emphasised the difficulty in determining the appropriate licensing version, especially as authorities combine and leverage the data in various ways, including the application of artificial intelligence technologies. The challenges posed by licensing further exemplify the complexities of open data and highlight the need to find solutions that facilitate easier access, promote interoperability and offer clarity in licensing terms, ultimately fostering a more inclusive and accessible open data ecosystem.

Technical barriers such as data silos, poor data quality and legacy technology-related issues, along with licensing challenges, hinder the full realisation of the potential of open data. Overcoming these challenges requires finding solutions that facilitate unrestricted access, promote interoperability and offer clarity in licensing terms. It is essential to strike a balance between openness and legal considerations, ensuring that licensing frameworks enable a more inclusive and accessible open data ecosystem. By doing so, public institutions can unlock the value of open data and facilitate collaboration, innovation and evidence-based decision-making. However, simply having fast and easy access to data is not enough for public institutions to create public value. Many of these barriers are intrinsically linked to non-technical barriers, such as legal restrictions, institutional processes and organisational culture.

3.2. Institutional barriers

Despite the increased awareness of data as an asset integral to policymaking, service delivery, organisational management and innovation, organisational and cultural factors still make data-driven value creation in public institutions difficult to achieve. This section of the summary paper will explore some of the institutional barriers to value creation from reuse of open data by public institutions, namely:

- the need to consider public institutions as data reusers from the outset of the data publication process;
- the lack of 'self-awareness' of public institutions that they are open data reusers;
- the lack of contact between the person or body responsible for the open data portal(s) and the ecosystem of users.

To create value out of open data, it is important to identify exactly what can be high-value and high-impact data for use and reuse. This implies a good understanding of the data users' needs and the interest and capacity in reusing data. To have a public service that is skilled in the recognition and use of data as a core component of effective governance is therefore crucial to the successful uptake of data use and reuse practices within public institutions ⁽⁶⁾. However, insights from the data.europa.eu campaign reveal that, despite the recognition of the relevance of data and of its potential to improve decision making and governance, public institutions are still undergoing a mindset change to see themselves as actual data users and reusers. When analysing the answers of the second mini-survey, it is striking that while the majority of respondents (92.52 %) reported considering data demand by other administrations important, they also indicated that their administrations do not deliberately target use from other entities (59.7 %).

⁽⁶⁾ Ubaldi, B. (2013), 'Open government data: Towards empirical analysis of open government data initiatives', OECD Working Papers on Public Governance, No. 22, OECD Publishing, Paris (<u>https://doi.org/10.1787/5k46bj4f03s7-en</u>).



The April 2023 focus group discussion further confirms this point. Many of the participating government data experts lamented that too often it is still difficult for public officials to grasp the concrete benefits of the data economy.

However, evidence from the campaign also shows signs of progress in this area, for instance the case of Estonia. There is more open data reuse in the public sector than is known and more than the actual reusers are aware of, as they often do not realise they are using open data. When detecting and sharing these cases, data reuse by public institutions can be further inspired and promoted. The reuse of open data seems under-documented and is not common practice for European data providers or ecosystem facilitators to measure public institutions' data demand. This area will be explored more in depth in the next section (<u>4. Empowering public institutions to reuse open data</u>). But their experiences also show how much still needs to be done in order to streamline this process across the EU and achieve an effective use of data that actually creates both economic and social value.

The lack of awareness in identifying valuable ways to reuse data is closely related. The lack of awareness in identifying such datasets is a common challenge faced in many countries. In order to fully realise the potential of open data, it is essential for public officials to be able to recognise the datasets that can be used and reused to generate value, and those that should not receive further resources. This requires a deep understanding of the potential uses of different datasets and knowledge of the specific needs of potential users who may be interested in accessing and reusing them.

In addition, the value of the data can be limited if public sector information cannot be reused, and data transparency may be impeded if data formats are difficult to access. Open data portals should allow the reuse of open data, ensure the efficiency of data transmission and enable professional initiatives based on data reuse. However, the creation of open data portals does not mean that the data they publish are prepared for professional reuse. As explained by Alberto Abella, data modelling expert and technical evangelist, during the focus group and expert interview, there are portals that exhibit professional apparent reusability ⁽⁷⁾ but do not allow reuse of their data; he defines such portals as 'pretender open data portals'. This can be understood as both a technical and a cultural barrier. Sometimes, the government's interest is in presenting data in a particular fashion to achieve political visibility, thereby limiting the use and provision of data to stakeholders who are interested in data reuse. This can be traced back to a lack of awareness

⁽⁷⁾ Abella, A., Ortiz-de-Urbina-Criado, M. and De-Pablos-Heredero, C. (2022), 'Criteria for the identification of ineffective open data portals: Pretender open data portals', Profesional de la información, Vol. 31, No 1, e310111 (<u>https://doi.org/10.3145/epi.2022.ene.11</u>).

of whether datasets for public reuse are of high value and to a lack of accurate measurements to assess such value and the impact of datasets. This issue was also highlighted by many of the data experts who took part in the focus group and webinar in the context of the data.europa.eu campaign. A representative of Opendata.fi presented a compelling example of this issue. They are using portal user statistics, such as the number of downloads of datasets, to assess the impact and relevance of their data. However, a careful analysis and research on the usage of such data revealed that the high number of downloads was due not to the high value of the datasets, but to the poor quality of the data. Individuals and organisations were downloading the data multiple times not because it was of high value, but because it was not maintained properly. This shows that analysing portal user statistics is not enough to measure and assess the demand of data reuse, they either rely heavily on portal user statistics or do not carry out this assessment at all.

Another similar issue highlighted during the focus group discussions is that governments may not consistently publish data or may shut down access to datasets without notifying users. This can cause problems for businesses that rely on the data for their operations or decision making. To address this, governments need to establish key datasets of 'high value' at both the EU and national levels. This requires careful consideration of the needs of different Member States, as they may have different priorities and use cases for the data. It is important for governments to establish clear policies and procedures for publishing and maintaining data, to ensure that it is accessible and reliable for all users.

Coordination or consistency issues may arise when the government seeks to impose consistency across the various rulemaking processes, data and portals that enable access to government data. Public agencies often have their own unique set of data, formats and standards, which can lead to unharmonised government data. This creates challenges for users trying to determine which piece of data is valid or trustworthy. **Accessibility is another critical aspect of open data and knowing the source of the data being searched for is crucial.** While a single open data portal may help to ensure integration and shared data input from various sectors of government, it is not always the best approach. However, when a decision is made to establish a single portal, it should be developed through a collaborative approach to ensure ownership and sustainability.

4. Empowering public institutions to reuse open data

However, regardless of the lack of explicitly targeted actions and available metrics to measure the use and reuse of open data, Member State authorities are still using these data to design policies and services for their citizens.

The actions needed to favour open data reuse by public institutions are not, per se, different from any well-designed and user-centric open data publication. Any purposeful, user-centric data publication effort would also be effective for the public sector as a data reuser.

What is important is to actually recognise the public sector as an important data reuser. Despite the scarce literature, existing research and the input gathered in the course of this campaign confirm that the public sector is one of the most important data reusers. The single most important action is therefore to avoid the rhetoric of open data as mostly a vehicle of economic development, which would lead to overlooking the public sector as a reuser. Simply removing this deeply rooted idea from the approach of open data publication is likely to reduce the bottlenecks.

Such awareness is particularly important today, due to events stemming from the COVID-19 pandemic. Data analytics has clearly become a core requirement for an effective public sector. Governments are required to radically improve their capacity to monitor and make sense of reality, and to anticipate forthcoming crises. In this context, open data reuse proved to be an important learning opportunity for building a government-wide data analysis culture. Thus, data reuse by the public sector becomes not just a demand that requires a response, but a strategic opportunity that needs to be actively fostered.

The question then becomes: how can the visible demand be met and the missing demand be stimulated? The discussions held during the webinar, the interviews and the focus groups point to four main areas of intervention, similar to those in place for stimulating reuse by any community.

4.1. Raising awareness

The need for greater visibility of public sector data reuse is particularly urgent because of the relative secondary relevance that it has achieved in the open data debate, where reuse by start-ups and civil society takes a prominent role. It is therefore particularly important, according to stakeholders surveyed and interviewed, to make an effort to give visibility to good practices across open data portals and dissemination activities. Many portals, in an effort to foster reuse, include a dedicated section on reuse cases; these should include cases of reuse by the public sector.

Similar 'mainstreaming' of public sector reuse should take place across dissemination and training activities, including the data.europa.eu academy, and in the impact assessment studies, to ensure that this impact is captured and visible.

4.2. Policy and regulatory frameworks

Fostering reuse by the public sector is part of a wider remit to make government more data-driven and evidencebased. Reuse will happen mostly because of an increasing quantity and quality of data demand by the public sector. Fostering this demand can happen directly through awareness-raising measures. But it mostly happens indirectly, as a result of greater demand for evidence-based policymaking, which emerges from the present research work as the main use case for open data use by the public sector. It is important that the work on open data publication is aligned and coordinated with efforts related to improved quality of policymaking and the new generation of 'predictive' policy tools. One example is the American Foundations for Evidence-Based Policymaking Act of 2018, which requires all agencies to spell out their data needs relating to planned regulatory efforts. The act has resulted in a healthy systematic exercise of data needs assessments by public agencies, which can then be systematically linked to efforts on the data publication side. Another example of a measure that favours the reuse of open data is the <u>EU open data</u> <u>directive</u>, which aims to promote the reuse of public sector information by making it more accessible and available. Importantly, the directive defines six categories of high-value datasets: geospatial, earth observation and environment, meteorological, statistics, companies and mobility. In light of this, in January 2023 the Commission published a <u>list of high-value datasets</u> that public sector bodies will have to make available for reuse, free of charge, within 16 months.

Only when data needs are assessed and visible will it be possible to pursue a 'publication with purpose' strategy.

4.3. Monitoring

There is a clear need for better monitoring of data reuse by the public sector, where there is currently a very visible gap. Based on the comparative analysis and the information from the campaign, it seems that the only available hard data point is for Spain and that this is based on a survey of data portal managers, not users.

It remains true that, since a fundamental open data principle is to reduce any friction to data reuse – such as the need for registration to an open data portal – tracking who is using data for what purpose remains elusive. There is a fundamental tension between the goals of frictionless reuse and purposeful publication. However, this should be treated is a trade-off, not an alternative. The absolute exclusion of any form of registration before accessing data should be reconsidered.

However, even without any form of registration, there are several possible methodologies to monitor reuse. Firstly, qualitative monitoring remains important. Simple dialogue with users is a necessary tool for open data publishers and should be actively pursued through one-to-one interviews and focus groups (either online or offline).

Surveys could be embedded in other promotion and community building activities. For instance, when there is a major event related to open data, surveys of participants should be carried out on a regular basis. Community building activities in particular could help create panels of respondents that could be surveyed over time to identify variations and track progress.

Open feedback surveys (through a feedback form on the website), while not providing representative samples because of self-selection effects, can still provide useful insight into how people use data and how data publication can be improved.

As a caveat, based on the lessons learnt in the present study, it is worth mentioning that there is a fundamental lack of awareness about data reuse by the public sector that hinders surveys, as many respondents struggle to understand the question. Activities to raise awareness are therefore also important to improve the quality of responses.

Lastly, partial monitoring could also be carried out through monitoring log access, in particular in cases where a public administration uses a single infrastructural network to access the internet. Logs could reveal the relative importance of access by public administration to the open data portals.

4.4. Data stewardship

There was consensus among experts involved in the focus group that the data steward role can be crucial to ensure reuse by the public sector – just as well as in the private sector.

Data stewards play a crucial role in open data portal management, as they are responsible for ensuring the accuracy, quality and usability of the data published on the portal. They act as custodians of the data, overseeing the entire data lifecycle from acquisition to publication and beyond, and are tasked with maintaining the integrity and security of the data. In addition, data stewards work closely with data providers to ensure that the data is properly formatted and appropriately licensed for reuse, and they may also collaborate with external partners such as developers and researchers to encourage the development of innovative applications and services based on the data. Moreover, data stewards can provide technical assistance and support to users of the portal, helping them to access, understand and use the data. This can include providing documentation, tutorials and other resources to facilitate data discovery and use. Overall, data stewards are important in open data portal management, as they help to ensure that the data published on the portal is of high quality, relevant and accessible to a broad range of users, ultimately promoting transparency, accountability and innovation. For instance, Slovenia is currently working with the OECD to codify the role of data steward.

There is still plenty of ground to cover: evidence from the survey shows, for instance, that only 40 % of public administrations deliberately target reuse by other administrations.

5. Conclusions and insights for data.europa.eu

The analysis of the data.europa.eu campaign reveals that there is still a long way to go for European public institutions to fully unlock the benefits of open data for the public sector. The problem lies in the fact that public institutions still struggle to recognise themselves as data (re)users and not just data holders. And when they do, there are limited metrics available to assess the demand of data and the impact of data reuse. As equally highlighted throughout the previous discussion paper, simply publishing datasets on a data portal will not generate data reuse or the associated benefits for public sector processes and results. The evidence gathered – especially through the webinar, the focus group and the expert interviews carried out within the data.europa.eu campaign – shows that European public institutions overwhelmingly recognise that open data is an important topic for evidence-based policymaking and data-driven public service delivery, but barely take action on it. The need identified by the OECD to 'move from hype to impact' is still very much present.

The implications of the actions listed above for data.europa.eu are quite straightforward.

When it comes to raising awareness and communication, any action should also contain examples of data reuse by the public sector. Gathering and communicating such examples and use cases greatly helps in understanding the importance of the role of the public sector as a data reuser.

When it comes to policy and regulation, it would be beneficial to align the 'better regulation' activities and roadmaps of the Commission with the open data publication activities, in order to better explore the internal data needs. Even more, it would be helpful to facilitate a similar alignment and data needs analysis for all European public administrations, for example by providing examples, best practices and methodologies on how to map data needs for policy and regulatory purposes.

Existing monitoring activities such as surveys should be revised to ensure that data reuse by the public sector is included. It would be useful to create a panel of users, based on the wide current community, that could be used for further surveys.

The figure of data stewards remains central to favour reuse, so examples, best practices and methodologies on the role of data stewards should be included in the support activities – not specifically for public sector reusers, but in general.

6. Bibliography

(1) Abella, A., Ortiz de Urbina Criado, O., De Pablos Heredero, C. and García Luna, D. (2022), 'Reusing open data in Spain III. October 2021.' Madrid: ESIC EDITORIAL.

(2) Abella, A., Ortiz-de-Urbina-Criado, M. and De-Pablos-Heredero, C. (2022), 'Criteria for the identification of ineffective open data portals: Pretender open data portals', *Profesional de la información*, Vol. 31, No 1, e310111 (<u>https://doi.org/10.3145/epi.2022.ene.11</u>).

(3)van Ooijen, C., da Silva Carvalho, N., Iordache, A. et al. (2022), 'Measuring data demand within the public sector: Discussion paper 2021', Publications Office of the European Union, Luxembourg (<u>https://data.europa.eu/doi/10.2830/54555</u>).

(4)Ubaldi, B. (2013), 'Open government data: Towards empirical analysis of open government data initiatives', OECD Publishing, OECD Working Papers on Public Governance, No 22, Paris (<u>https://doi.org/10.1787/5k46bj4f03s7-en</u>).

(5) van Ooijen, C., Ubaldi, B. and Welby, B. (2019), 'A data-driven public sector: Enabling the strategic use of data for productive, inclusive and trustworthy governance', OECD Publishing, OECD Working Papers on Public Governance, No 33, Paris (<u>https://doi.org/10.1787/09ab162c-en</u>).

(6) Vasconcelos, J.R., Kim, Y., Mocan, S., Martins, O.I.J., Halsema, R.A. and van Ooijen, C.W. (2023), 'Interoperability – Towards a data-driven public sector (English)', World Bank Group, Equitable Growth, Finance and Institutions Insight, Washington, D.C. (<u>https://doi.org/10.1596/38520</u>).

7. Annex: Methodology

The summary paper relies on the material gathered through the campaign and complements this with expert interviews.

The data.europa.eu campaign consisted of a scoping paper, a webinar, two mini-surveys, a blog series and a focus group, which provided valuable insights into the challenges and opportunities of data-driven value creation in the public sector. The mini-surveys and the focus group were conducted among public institutions across Europe, while the webinar and blog series targeted a wider audience of data professionals and enthusiasts. These activities offered valuable insights into the current state of open data reuse practices among public institutions in Europe.



7.1. Mini-survey structure

The mini-surveys of the campaign, while not being statistically representative, provided new data points that supply a useful snapshot of the perception of the community. The purpose of the surveys was to improve understanding of the demand for open data among public administrations in Europe, along with the challenges they face in adopting and implementing open data policies. We surveyed a diverse range of public administrations across Europe, including national, regional and local governments and public institutions and agencies, and we collected a total of 108 answers to the first survey and 548 answers to the second one.

7.1.1. Mini-survey n1

- For which organisation do you work?
- Had you given thought to data reusers from public institutions prior to this webinar? For instance, a Member State reusing data from Eurostat or a municipality reusing statistical data published.
- If yes, please tell us what issues/questions have come up in your organisation in relation to this data reuser group.
- Has your organisation undertaken activities to assess the demand of data reusers from the public sector?
- If yes, what are your organisation's experiences with data reusers from the public sector so far? In what way are they different from other reuser groups?
- Has your organisation ever acted as an open data reuser itself (e.g. using data published by other EU institutions or public authorities in Member States)?
- If yes, please tell us which data you used, for what purposes and what led you to reuse these data.

7.1.2. Mini-survey n2

- For which organisation do you work?
- What is the name of the organisation?
- In which country do you work?
- Does your administration use open data from other government entities for internal use or for providing new data services?
- If yes, can you mention some examples?
- Considering the activity of your organisation as a publisher of data, how important would you consider that is the demand by other administrations (from 0 to 5, with 0 being 'Not important' and 5 being 'Very important')?
- Does your administration deliberately target use and reuse of open data by other administrations?
- If yes, can you please explain how?
- Does your administration regularly measure demand and reuse? What indicators do you use?

7.2. Webinar: data demand and reuse in the public sector

The data.europa.eu academy on Data Demand and Reuse in the Public Sector brought together representatives of EU institutions and Member States with basic knowledge of open data and interest in data reuse in the public sector. No prior knowledge was required.

Learning objectives

- Gain insight into the value of data reuse within the public sector.
- Learn about different approaches to measure public institutions' data demand.
- Strengthen the understanding of a demand-driven approach to data publication.
- Reflect on how data.europa.eu can support data providers in taking on board public institutions' data demand.

The aim of the webinar was to provide an introduction to data reuse by public institutions and the importance of engaging with and measuring data demand by this specific group of users. Various expert speakers explained how they engage with data reusers from public institutions, measure their data demand and incorporate the insights in their open data policies.

Topics covered

- What is public institutions' data demand, why does it matter and how can it be measured?
- What are best practices in engaging with data reusers from the public sector and measuring their demand?
- How can data.europa.eu support potential data providers?
 - community-building support;
 - \circ $\;$ indicators to measure public institutions' data demand;
 - other portal features.

Expert speakers

- Frederika Welle Donker, researcher at the Knowledge Centre Open Data of Delft University of Technology, the Netherlands.
- Antonin Garrone, product owner at Etalab, the French task force for data policy.

7.3. Focus group: how open data is used by the public sector

The data.europa.eu academy focus group 'How open data is used by the public sector' brought together a group of interviewees, from EU institutions, national public administrations and regional/local authorities, who are interested in exploring the ways in which data can be used to create value in the public sector.

The focus group facilitated a discussion around open data demand and reuse by the public sector. Participants, and particularly panellists, were encouraged to share their experiences, insights and best practices, along with their ideas for how to overcome barriers to success in this area.

Learning objectives

- Identify methods and good practices in assessing public institutions' data demand.
- Advance government-to-government open data reuse by sharing knowledge and collaborating.
- Understand the needs of open data reusers from public sector institutions in Europe.
- Learn to create the support mechanisms public administrations require for data reuse, such as data literacy programmes and data reuse partnerships.

Key questions

- Are you using data from other public administrations? How?
- Can you share examples of impactful open data reuse by public institutions and tell us why these are compelling?
- What have you learnt from these success stories, or the lack thereof, on what is needed to turn public servants into data reusers?
- What are the obstacles your organisation faces relating to data reuse?
- How do you think data.europa.eu can support data sharing and data reuse across public administrations?
- What advice can you give on promoting the reuse of open data for public policy making and service delivery, and approaching the relevant reusers?

The participants focused on these questions but also held an open discussion. These interactions, particularly among the panellists, produced very valuable results.

7.4. Expert interviews structure

The first expert interview was conducted with data modelling expert and technical evangelist Alberto Abella, who offered insights into measuring data demand and reuse within public institutions. The second interview was conducted with Stefaan Verhulst, co-founder and chief research and development officer of the Governance Laboratory at New York University (GovLab), who provided insights into the governance challenges of data-driven governance in the public sector.

Alberto Abella, data modelling expert and technical evangelist, FIWARE:

- What are some of the key indicators that public institutions should use to measure the demand for data within their organisations?
- In your opinion, what are some of the key benefits of measuring data demand and reuse for public institutions?
- What are some of the challenges that public institutions may face when trying to measure data demand and reuse, and how can they overcome these challenges?
- How can data modelling support the measurement of data demand and reuse within public institutions?

Stefaan Verhulst, co-founder and chief research and development officer of the Governance Laboratory at New York University (GovLab):

- In your experience, what are some of the key governance challenges that public institutions face in promoting data-driven governance?
- How can public institutions foster greater collaboration and partnership in order to enable data-driven governance?
- What are some of the challenges that public institutions may face in promoting transparency and public trust in data-driven governance, and how can they overcome these challenges?
- How can public institutions ensure that they are effectively measuring the impact of data-driven governance on the delivery of public services, and what are some of the key indicators that they should be tracking?

ISBN: 978-92-78-43685-8

