European data spaces and the role of data.europa.eu

This report has been prepared as part of data.europa.eu, which 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 the following.

DISCLAIMER

Contents

1.   Introduction

In February 2020, the European Commission published its European data strategy (see section on Legislative landscape) (1). The strategy aims to create a single market for data that freely flows within the European Union (EU) for the benefit of citizens, businesses, researchers and public administrations. At the core of this strategy are common European data spaces covering 10 strategic fields. By providing relevant data infrastructure and governance frameworks, these data spaces should facilitate the pooling and sharing of data within and across sectors in Europe.

Data sharing and the initiation of common European data spaces are supported by policy, legislative and funding measures from the Commission. This changing data sharing landscape affects the role of data.europa.eu as an official portal for European open data. An online panel was held on 7 June 2023 to discuss the potential of data spaces for Europe's citizens, businesses and economy. This report captures the outcomes of the panel discussion and presents considerations for data.europa.eu in light of an evolving data sharing landscape.

As background, basic concepts of data spaces and relevant legislation of Europe's data economy are described. Then, the report presents the outcomes of the panel discussion and offers suggestions for data.europa.eu and the data spaces community. Specifically, the public procurement data space (PPDS) is showcased as an example of a data space initiative in practice. Similarly, the Data Spaces Support Centre (DSCC) is presented as an example of the supporting structures data spaces require to be successful. Digital literacy is also discussed as an important enabler of Europe's digital ambitions.

2.   The concept of data spaces

In this section we summarise a few basic concepts that are useful for understanding the subsequent parts of the report. Data spaces are a novel topic and the literature available for reference is limited. Consensus on terminology and definitions is also not yet achieved between the several sources active in observing and describing the phenomena. In the following section, the authors attempt to present the topic according to the EU canon, namely EU regulations and other official EU documents, along with research produced or sponsored by the EU, such as from the Horizon Europe and digital Europe programmes.

2.1. What is a 'data space'?

According to the Commission staff working document on common European data spaces (2), a data space 'brings together relevant data infrastructures and governance frameworks in order to facilitate data pooling and sharing'. Another definition is provided by the DSCC glossary (3), namely 'an infrastructure that enables data transactions between different data ecosystem parties based on the governance framework of that data space'. Furthermore, data spaces must be 'generic enough to support the implementation of multiple use cases'. A data ecosystem, in turn, is 'a loosely coupled set of autonomous parties engaging in data sharing.'

---

(1)   https://digital-strategy.ec.europa.eu/en/policies/strategy-data.
(2)   https://digital-strategy.ec.europa.eu/en/library/staff-working-document-data-spaces.
(3)   Version 1.0 at
        https://dataspacessupportcentre.refined.site/space/Glossary/55443460/DSSC+Glossary+ %7C+Version+1.0+ %7C+March+2023?atta
        chment=/rest/api/content/55443460/child/attachment/att110362680/download&type=application/pdf&filename=DSSC-Data-
        Spaces-Glossary-v1.0.pdf.

It is useful to highlight that a data space is not just an organisation or a technology platform, nor is it only data or services being exchanged under a governance model. Rather, a data space is the collective workings of all those elements together. As an analogy, think of a murmuration of starlings that decide to fly in patterns without being strictly 'organised', orchestrated by a central entity or using an external shared artifact (a 'platform') to regulate their behaviour and tell them where to fly.

A murmuration of starlings (3), https://flic.kr/p/8Ve8tA.

Nonetheless, the term data space is commonly used as a synonym to refer to the organisation that represents the data space's participants and provides governance to the space. This organisation is sometimes called the '**data space *governance authority***'. For example, the Europeana Foundation ([4]) operates the cultural heritage data space ([5]) for the EU and is its 'data space governance authority'. The data space governance authority may change over the life of a data space, like all other services provided to participants. The data space itself is the network of participants and their interactions in sharing data and data-related services.

Instead, the term '**data space *initiative***' is used to differentiate an actual data space from every initiative that is not yet a data space, but is aimed at paving the way for a future data space. Most of the data space coordinated support actions funded by the digital Europe programme ([6]) are of this kind.

The EU calls '**common European data spaces**' all data spaces and data spaces initiatives funded by the EU, each typically focused on addressing the needs of one sector (agriculture, Green Deal, etc., see Section **Error! Reference source not found.**). They are 'common' because they are aimed at serving the whole of the EU, without restrictions to participation. As such, these data spaces are commonly referred to as being 'open'. At the same time, the common European data spaces are not intended to be exclusive; other data spaces exist and can emerge in the future, serving the same sectors ([7]).

---

[4]   https://pro.europeana.eu/about-us/foundation.
[5]   https://www.europeana.eu/en.
[6]   https://digital-strategy.ec.europa.eu/en/activities/digital-programme.
[7]   For example, in the summer of 2023 the government of the Netherlands launched a Green Deal data space with a specific focus on the environmental impact of non-residential buildings: the 'Datastelsel Verduurzaming Utiliteit'. Read about it at https://ishare.eu/inspiration/smart-energy-dsweu/.

The common European data spaces together form a '**common European data space**' (singular). This alludes to the concept of a 'European common market' ([8]) and to how participants can perform their activities across all data spaces under a similar governance and with the least friction possible.

2.2. Why are data space service-focused and user-centric?

Data spaces can be seen as public squares where a market takes place. In a traditional market, the citizens (data consumers/beneficiaries) establish an authority that regulates and administers the market *in their name*, typically the municipality (the governance authority). Each farmer (data rights holders/providers), for example, can participate and come to the market (data space) and offer products (data) as long as they follow the rules that the citizens have agreed upon and the municipality enforces in their name ([9]). The municipality will also enforce the basics of social interaction, such as preventing theft or antisocial behaviour.

Similarly, in a data space, participants establish an authority that represents them (the data space governance authority) and participants can offer data and/or data related services (e.g. storage, processing, AI functionality) or consume the data and data services offered by others. Of course, a few participants will specialise in offering services, while others will use the data space just to purchase them, as in most markets. As in the case of a public square, it is rare that access is not open to everybody, as long as they follow the rules. However, the rules may include limiting factors, such as a requirement for membership. The EU-funded data spaces are intended to be open to anybody.

The services are the protagonist of this model. The data space governance authority, by contrast, may not provide any data services, except for the operations of the data space's governance. Because of this service focus, satisfying the participants by providing them with the data and services they need is not the responsibility of the data space governance authority, but rather of the other participants. This is sometimes referred to as the user-centricity of this model. **The users are ultimately the ones who decide if there is value for them participating in a data space, and the data space is expected to naturally evolve according to their needs, expectations and perspectives**.

2.3. Why are data spaces decentralised?

Data spaces aim at being **decentralised**. The Horizon 2020 'Open DEI' ([10]) project defines 'decentralisation' as one of the four design principles of data spaces, together with 'data sovereignty', 'data level playing field' and 'public-private governance' ([11]). Decentralisation refers to the distribution of power, authority or decision-making away from a single central entity or authority, instead dispersing it among multiple participants. By promoting distribution and reducing reliance on a central point of control, decentralisation aims to enhance transparency, resilience and democratisation in various systems and applications.

---

[8]   https://european-union.europa.eu/priorities-and-actions/actions-topic/single-market_en.
[9]   By the way, nobody stops farmers from buying food from other farmers; they are citizens too.
[10]   Open DEI received funding from the European Horizon 2020 programme for research, technological development and demonstration under grant agreement No 857065, https://www.opendei.eu.
[11]   https://design-principles-for-data-spaces.org.

The picnic table analogy to understand decentralisation

We can use the 'picnic table analogy' to explain the basic difference between centralised and decentralised data spaces.

A data space is like a picnic among friends.

In a <u>centralised</u> picnic, a few friends decide to meet at the park to have a picnic. The friend at the centre of the picture hosts the picnic and has brought the table (the 'platform'). The other friends can offer consumable items such as food and drinks (the 'data') and equipment such as cutlery and crockery (the 'data services') that they put on the table. The friends can then examine what is available on the (central) table and drink and eat what they like with the cutlery and crockery that suit them.

The picnic takes place within the rules of the park (e.g. dogs are not allowed), the host and the friends. The host administers the table and governs the sharing of the food, drink, cutlery, etc., according to the rules set by the friends. The host can add her own rules.
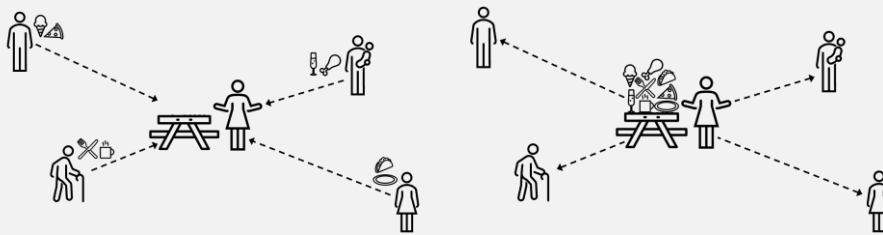
In a <u>decentralised</u> picnic, there is no single host. Everybody brings and takes whatever they feel like. Someone also brings a table. Anybody is free to come and go as they please. The picnic lasts as long as someone is still around. Each exchange is agreed between the involved parties. The rules of the park still apply.

The data space is not simply the sum of the tables and baskets. *The data space is the picnic*: the fact that 'something is happening' thanks to the participants who have come together to share, according to the rules of the park and any other rules the friends have agreed between themselves. There is no all-powerful host.
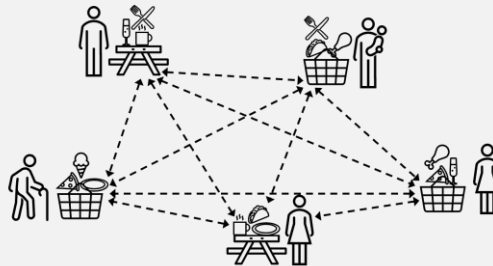
The picnic table, cups, napkins and cutlery are the 'soft infrastructure', i.e. the services that enable the exchange and consumption of drink and food.

In the centralised picnic, the downsides and dependency on the host are evident. The host may, unfortunately, offer a table that is wobbly, cups that are unsanitary or forget the cutlery at home. Their friends do not have much choice. If the host needs to go home, the table is gone and the picnic is over.

In a decentralised picnic, it may, unfortunately, happen that nobody has brought something necessary – nobody brought the napkins, say – but it is not likely, particularly if we have lots of friends. It may take longer for a decentralised picnic to reach a point of

balance where enough friends have joined, are satisfied, are staying and everything is available. The overall result is, however, a fairer and more 'resilient' picnic.

***

The reader may wonder where data.europa.eu would fit into this analogy.
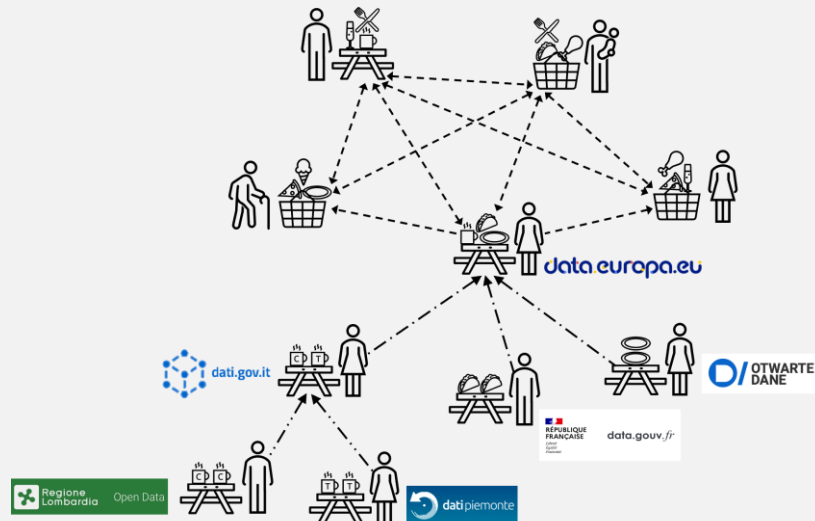
By operating a website open to anybody, data.europa.eu – like most open data portals – offers its services to anybody, including any data space participant of any data space. data.europa.eu may decide not to re-distribute the data that the data space participants might offer, in which case the data exchange would be in one direction from data.europa.eu to data spaces.

The data held by data.europa.eu is sourced from its own data providers: supranational, national, regional and local public administration in Europe. In the picture, the national open data portals of France, Italy and Poland are illustrated as examples.

In turn, the data providers for data.europa.eu source data from their own data providers. Each EU Member State sources data from its own regions and cities. In the picture, the Italian open data portal sources are the Lombardy and Piedmont regions (and all other regions that are not illustrated for convenience).

The various (dashed) lines connecting the participants show how different interoperability arrangements could be set up, such as different standards for describing the catalogues.

Being decentralised is not yet a necessary condition for classification as a data space in the EU. It should be noted that very few such systems have presently achieved that level of maturity in their technology and governance models.

A data space can be centralised. Centralised data spaces are most often based on a centralised platform offered by a data intermediator to the data sharing participants. This model is common today and available commercially ([12]). A few companies are also known to host their own centralised data space to, for example, incentivise and support partners to collaborate and innovate using the

---

[12] Among the commercial services known to the authors through the Capgemini Invent network are Dawex https://www.dawex.com/en/, Harbr https://www.harbrdata.com and the data sharing and collaboration services of Snowflake https://www.snowflake.com/en/.

company's data ([13]). In the centralised data space model, however, there is a high risk of the platform acquiring unfair power and advantage over the participants, possibly recreating the same issues we can see with data monopolies ([14]).

In the meantime, the Data Governance Act (DGA) (see section on Legislative landscape) is already used to prevent the issues of centralisation, by regulating data intermediaries and requiring structural separation between the data intermediation service and any other services provided. Under the DGA, for example, data intermediaries can charge for their services but cannot monetise the data (e.g. by selling it to another company or using it to develop their own products based on this data) and have to comply with strict requirements to ensure neutrality and avoid conflicts of interest ([15]).

2.4. Why are data spaces automated and have common standards?

Data spaces are designed to be extremely scalable. As a natural consequence, data spaces aim at being heavily (if not fully) **automated**. Scalability requires processes and procedures to be automated as much as possible and limits the need for human intervention. Technical standards are being developed to achieve this. Such **common standards** are important for automation, but also for interoperability between the numerous participants in a data space.

2.5. What is a 'data portal'?

In EU policymaking, a 'data portal' refers to websites set up by public administration entities, at any level of government, that act as catalogues to support the **discovery** of public data resources. Making the resources 'discoverable' refers to the whole set of functionalities offered to re-users that enable them to find out if the data they need exists and if so, to search for those datasets that best fit their needs. Data portals are generally aimed at being used by the public at large: citizens, businesses, researchers and organisations in general.

In many cases data portals act as *meta*-catalogues, meaning that they aim at making discoverable in a consistent and coherent way public data resources offered by other, related government entities, each of which may have their own data portal. Traditionally, in the EU, data portals are organised according to a hierarchy with data.europa.eu at the top, as it is a meta-catalogue of national data portals that in turn are meta-catalogues of their ministries, regions, cities, etc. ([16]). This is just a convention and is not enforced in legislation, for example in the open data directive (see section on Legislative landscape).

Besides discoverability, data portals can also be the **publication and storage** venues of such resources. As the use of real-time data and application programming interfaces (APIs) become more widespread, it is common for publication to happen as close to the source of the data as possible ([17]).

Narrowly speaking, data portals are technical solutions to provide discoverability, publication and storage functionality to users. However, providers of data portals typically run a **portfolio of initiatives by which the supply and re-use of public data resources are promoted**, of which maintaining the technical portal is just one such initiative. From their beginning, major portals such as data.europa.eu

---

([13])   For example, the Airbus Skywise service https://aircraft.airbus.com/en/services/enhance/skywise.
([14])   Anecdotally, among data space practitioners it is uncommon to hear the word 'platform', precisely in order to avoid misunderstanding and appearing to support the idea of centralisation.
([15])   https://digital-strategy.ec.europa.eu/en/policies/data-governance-act-explained#ecl-inpage-l4ihlqt9.
([16])   For example, the city of Barcelona will operate the data portal for Barcelona's public data resources, which will also be discoverable on Catalonia's data portal, Spain's data portal and data.europa.eu.
([17])   For example, Barcelona's public transport timetables are more likely to be published on Barcelona's public transport website than on the Spanish data portal.

have operated significant initiatives to inform, support and educate data providers and data re-users in addition to hosting the website. It is common, then, for the term 'data portal' to be used broadly to refer to the wider vision by which one institution promotes some set of public data resources coming from different public administrations.

2.6. What is the difference between a data space and a data portal?

If we use 'data portal' to refer to the discovery and sometimes publication and storage of data resources, then **a data portal is a centralised infrastructure offering a subset of the *technical* functionalities that a data space requires**.

In other words, a data portal could be seen as a type of participant in a data space whose objective is to make discoverable and/or available the data in which it specialises. Occasionally, it may provide other services for that same data, such as visualisation. That is what data.europa.eu does. Most of its services, as a portal, are focused on making EU data *discoverable*, and it points users to access the data from a separate source (e.g. the original data provider, such as a ministry in Italy or a city in Spain). In a few cases, typically for EU institutions, the data is accessible from data.europa.eu itself.

There is no universally recognised specification of the full set of functionalities of a data space. However, the Joint Research Centre, for example, lists 'publication and discovery' and 'storage' – functionalities typical of data portals – as just two of the technical functionalities expected of data spaces. Other functionalities listed for data spaces by the Joint Research Centre include 'transfer and exchange', 'identity, authentication and access control', 'interoperability', 'processing and analytics' and 'pooling and collaboration' ([18]).

If we refer to a data portal in the wider connotation as a programme aimed at promoting the supply and re-use of data resources, then the similarity with data spaces becomes stronger. Like data portals, data spaces are also aimed at fostering the supply and re-use of data resources. However, in most cases, the following apply.

(a) **The data resources in data spaces are most often not public:** they may include personal or confidential data, or data that is covered by strict intellectual property and licensing or contractual terms. (As many national data portals prepare to become the 'single access points' specified by the Data Governance Act to also make discoverable non-open datasets controlled by the government ([19]), this distinction may soon become irrelevant.).

(b) **The data resources in data spaces are not necessarily offered by public administration institutions,** but potentially by anybody, including individuals.

(c) **The data spaces are not necessarily operated as public services,** but most commonly as commercial services offered to their participants. Of course, the EU-funded data spaces will be operated as public services.

([18]) See Section 2.2 'Technical perspective' in European Commission, Joint Research Centre, Farrell, E., Minghini, M., Kotsev, A., Soler Garrido, J., Tapsall, B., Micheli, M., Posada Sanchez, M., Signorelli, S., Tartaro, A., Bernal Cereceda, J., Vespe, M., Di Leo, M., Carballa Smichowski, B., Smith, R., Schade, S., Pogorzelska, K., Gabrielli, L. and De Marchi, D., *European Data Spaces – Scientific insights into data sharing and utilisation at scale*, Publications Office of the European Union, Luxembourg, 2023, https://data.europa.eu/doi/10.2760/400188.

([19]) https://data.europa.eu/en/news-events/news/european-single-access-point-harvesting-guidelines-member-states.

3. Legislative landscape for data

The Commission envisions a data-driven society for the EU, where access to data and the ability to use it are key for innovation and growth. The value of the data economy in the EU is projected to reach EUR 829 billion in 2025, a 2.75-factor increase from EUR 301 billion in 2018 [20]. The EU is taking action to build this data economy, and with it, the European way of handling data. In this respect, the Publications Office of the European Union (which manages data.europa.eu) is supporting the implementation of data policies to move in this direction.

3.1. European data strategy

The European data strategy (February 2020) [21] is the foundation of work to make the EU a leader in a data-driven society and ensure that Europe's digital transition happens according to European values and principles. The data strategy sets out the vision of creating a single market for data, where data can flow freely within the EU and across sectors. This market should give access to high-quality data to various stakeholders, including businesses, researchers and public administrations. Within this context, existing rules on topics such as data sharing, privacy and competition are relevant and must also be respected. The panel discussed that various kinds of instruments are needed to achieve the goal of a single market for data, including policy, funding and legislative measures. In terms of legislation, the open data directive [22], the Data Governance Act [23], the Data Act [24] and the AI Act [25] are key legislative instruments supporting the data strategy. These legislative instruments are described below (including their relation to data, if data is not the primary topic of the legislation) and reflect the legislative measures referred to in the panel discussion. Furthermore, the relationship between the legislation and data spaces is described. The regulations described in this report are not intended to be an exhaustive list.

3.2. Open data directive

The directive on open data and the re-use of public sector information (entered into force in July 2019 and transposed into national law by July 2021) [22] 'provides common rules for a European market for government-held data'. This directive replaces the public sector information directive [26]. One of the main aims of the open data directive is to make public sector data (e.g. from public sector bodies in EU countries, at the national, regional and local levels) and publicly funded data (e.g. from meteorological institutes) reusable. The directive also mandated an implementing act of high-value datasets (HVDs), which followed on December 2022 [27]. HVDs are certain public sector datasets that have been identified as having important benefits for society, the environment and the economy across six categories: geospatial, earth observation and environment, meteorological, statistics, companies and mobility. HVDs must be made available for re-use with minimal legal and technical restrictions and be

---

[20] https://www.europarl.europa.eu/news/en/headlines/society/20210218STO98124/european-strategy-for-data-what-meps-want
[21] https://digital-strategy.ec.europa.eu/en/policies/strategy-data and https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX %3A52020DC0066.
[22] https://digital-strategy.ec.europa.eu/en/policies/legislation-open-data and https://eur-lex.europa.eu/legal-content/EN/TXT/?qid=1561563110433&uri=CELEX:32019L1024.
[23] https://digital-strategy.ec.europa.eu/en/policies/data-governance-act and https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX %3A32022R0868.
[24] https://digital-strategy.ec.europa.eu/en/policies/data-act and https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=COM %3A2022 %3A68 %3AFIN.
[25] https://digital-strategy.ec.europa.eu/en/policies/european-approach-artificial-intelligence and https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A52021PC0206
[26] https://eur-lex.europa.eu/legal-content/en/ALL/?uri=CELEX:32003L0098.
[27] https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=uriserv:OJ.L_.2023.019.01.0043.01.ENG.

free of charge. Overall, the directive is built around the concepts of transparency and fair competition in the internal market.

How the legislation relates to data spaces

Measures of the legislation	Connection to data spaces
— Stimulates the publication of dynamic data and the uptake of APIs — Limits the exceptions to charging above the marginal costs of dissemination for public data resources — Enlarges the scope of data covered by the original directive — Strengthens transparency requirements to avoid exclusive arrangements between the public and private sectors	The connection between the open data directive and the data spaces programme is loose. It is however useful to highlight a few points — The open data provided by public institutions is instrumental to most data applications, including the ones that will source data from data spaces — It is not necessary for the open data portal to 'explicitly' join data spaces, as they are – by definition – already available to anyone — Although most data spaces focus on protected data and the private sector, their output is not necessarily protected data. The collaboration of participants can in fact create datasets suitable for publishing under an open licence ([28])

3.3. Data Governance Act

The DGA (entered into force in June 2022 and applicable from September 2023) ([23]) aims to 'increase trust in data sharing, strengthen mechanisms to increase data availability and overcome technical obstacles to the re-use of data'. The panel summarised that the goal of the act is to boost voluntary data sharing.

How the legislation relates to data spaces

Measures of the legislation	Connection to data spaces
— Sets rules and means for the re-use of sensitive data held by public sector bodies — Creates the foundations for data altruism mechanisms — Provides safeguards for access requests from third countries in the context of non-personal data	Data spaces offer technical functionality and governance to establish rules for complex data sharing transactions, offering the means to share sensitive data in a trusted and secure manner
— Regulates data intermediation services to ensure providers function as trustworthy organisers of data sharing or pooling	A data space's rules must comply, inherit and can extend what is specified by the applicable legislation of the territory in which they

---

([28]) Depending on the licensing of the open data being re-used as input, for example under share-alike terms, participants may actually be forced to publish the derived data under similar licensing to be compliant.

	operate. In the EU, for example, the DGA specifically regulates data intermediaries
	Data space technology can enforce several of these rules
— Establishes the European Data Innovation Board (EDIB) to facilitate the sharing of best practices (particularly on data intermediation, data altruism, the use of public data that cannot be made available as open data, and the prioritisation of cross-sectoral interoperability standards)	When and if adopting recommendations from the DSCC regarding which 'building blocks' are preferred to operate EU-funded data spaces, the EDIB will make them de facto standards for Europe

3.4.   Data Act

The Data Act (DA)([24]) (adopted by the Commission in February 2022, granted political agreement in June 2023 ([29]) and currently undergoing formal approval) would introduce new access and use rights on data to make more data available for use in line with EU rules and values. Significant amounts of industrial data remain unused. The DA intends to boost the EU data economy by unlocking this industrial data. The EDIB, established under the DGA, will also advise and assist the Commission in developing guidelines and assessing the progress of DA initiatives.

How the legislation relates to data spaces

Measures of the legislation	Connection to data spaces
— Increases legal certainty for companies and consumers that generate data on who can use what such data and under which conditions, and incentives for manufacturers to continue investing in high-quality data generation	Data space technology aims at offering as much as possible the automated enforcement of both (a) the legal framework in which the data space operates, and (b) the rules specific to that data space that form its governance
— Prevents abuse of contractual imbalances that hinder fair data sharing	The rules enforced in data spaces can be used for the empowerment of data users and data holders and help establish a healthy balance between the rights and interests of any stakeholder involved by enforcing mutually acceptable rules in a data transaction
— Provides means for public sector bodies to access and use data held by the private sector that is necessary for specific public interest purposes	The DA does not specify the exact data sharing model by which public sector bodies will access private sector data. Data spaces, however, may be a natural model to do that when they become mainstream, rather than ad-hoc arrangements. It is easy to imagine a public sector body joining the data

---

([29])   https://ec.europa.eu/commission/presscorner/detail/en/ip_23_3491.

	space(s) holding the data that it requires, and using its/their pre-existing functionality to gain access to the data
— Sets the framework conditions for customers to effectively switch between different providers of data-processing services	The use of standards and the decentralised characteristic of data spaces maximise the options and flexibility of participants in switching data-processing service providers

3.5. AI Act

The artificial intelligence (AI) Act ([25]) (proposed in April 2021 and currently being negotiated in the European Parliament ([30]), aims to boost AI research and industrial capacity in Europe while ensuring safety and fundamental human rights. In essence, AI systems must be analysed and classified according to the risk they pose to users ([31]). Certain obligations are established for providers and users depending on the specific uses of AI. For example, certain high-risk practices such as biometric surveillance are prohibited and certain general-purpose AI systems such as (generative) foundation models must implement transparency measures ([32]). Furthermore, exemptions are considered for research activities and AI components provided under open-source licenses, in order to boost AI innovation and support small and medium-sized enterprises.

How the legislation relates to data

Building high-performance and robust AI systems relies on high-quality data. This requirement for high-quality data creates links with other initiatives. For example, the DGA aims to provide the right infrastructure for building such systems. Moreover, open data and data spaces are sources of large and diverse datasets for AI systems ([33]).

How the legislation relates to data spaces

Measures of the legislation	Connection to data spaces
— Establishes risk-management rules for AI systems — Sets transparency rules for AI systems	Data spaces can enforce standards on metadata quality, such as information on data provenance and lineage, which can be used as part of a risk assessment and for transparency reporting
— Requires that summaries of the copyrighted data used for the training of AI systems are made publicly available	Data spaces can enable the necessary due diligence to ensure that the information regarding data used for training the systems is transparent and that the reuse of copyrighted data is legitimate

([30])   https://www.europarl.europa.eu/news/en/press-room/20230609IPR96212/meps-ready-to-negotiate-first-ever-rules-for-safe-and-transparent-ai.

([31])   https://www.europarl.europa.eu/news/en/headlines/society/20230601STO93804/eu-ai-act-first-regulation-on-artificial-intelligence.

([32])   https://www.europarl.europa.eu/news/en/press-room/20230505IPR84904/ai-act-a-step-closer-to-the-first-rules-on-artificial-intelligence.

([33])   https://data.europa.eu/en/publications/datastories/open-data-and-ai-symbiotic-relationship-progress.

4. Common European data spaces

The concept of common European data spaces is to 'bring together relevant data infrastructures and governance frameworks in order to facilitate data pooling and sharing' (34). The objective is to create 'EU-wide common, interoperable data spaces in strategic sectors that overcome existing legal and technical barriers to data sharing'. These data spaces together will form a single market for data in Europe that allows more data to be made available for use in the economy and society. The intention is to make data exchanges are trustworthy and secure. The panel summarised this objective as freeing data from silos to stimulate the European data economy.

The anticipated outcome of reducing legal and technical barriers to data sharing is that new data exchanges are enabled for different stakeholders. In this context, data can be used for new purposes than for which it was originally collected. This new use of data allows the creation of more value through new services, products, better policymaking and better public service delivery.

Most common European data spaces are sector-centric. Ten sector-focused common European data spaces were originally announced in the European data strategy (Figure 1). Since then, additional data spaces have been created, including for tourism, cultural heritage, media and language. The panel described that the Commission favours a bottom-up approach that is community and stakeholder-based for the development of data spaces (see section on PPDS). In this approach, each sector or domain community must engage and assess issues such as whether data sharing is not happening in their space. The community should then try and solve the issue, including by setting up the right governance for the data space.

The panel emphasised that a data space is not solely a technology platform or tool like a data lake. A data space is also not solely a catalogue of data, nor a technical infrastructure that allows authorities to exchange data. According to the staff working document on common European data spaces, these data spaces are guided by several design principles about data control, governance, respect for EU rules and values, technical data infrastructure, interconnection and interoperability, and openness.
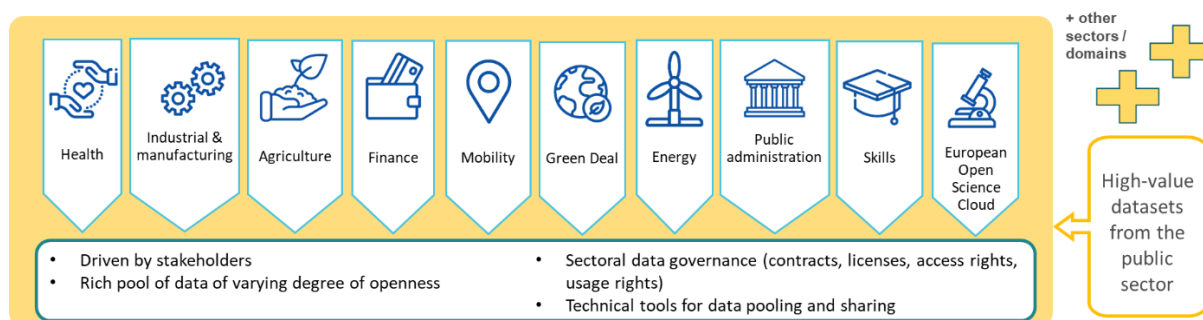
Figure 1: **Overview of the common European data spaces**

Through the common European data spaces, the Commission intends to harmonise the development of data governance in all sectors. The panel described that the legislation emerging under the European data strategy, such as the Data Governance Act, Data Act and open data policy, is building a horizontal legislative framework that will play a key role in developing the common European data spaces. For example, the DA considers the common European data spaces through an interoperability provision. The panel highlighted that interoperability between the various sector-specific data spaces

---

(34) https://digital-strategy.ec.europa.eu/en/library/staff-working-document-data-spaces.

is important because it enables data to flow across the different sectors. This in turn allows the European market for data to be fully deployed and operational.

The Commission also provides various funding measures, some of which are linked to the digital Europe programme, to support the development of data spaces. The Commission additionally supports the development of common European data spaces through initiatives and assets for stakeholders. A hallmark implementation action is the DSSC, which provides tools and creates knowledge for all data spaces. These tools and knowledge should enable data spaces in Europe to develop in a coordinated manner (see section on the DSCC). The Commission is also working on a technical solution called Simpl ([35]), which will be both an open-source middleware (Simpl Open) and a software as a service (Simpl Live), and will be available to all EU-funded data spaces in order to operate their services.

Given the range of policy, legislative and funding measures from the Commission for data sharing, the panel concluded that all of these instruments will be implemented in the common European data spaces. Stakeholders in the data spaces are then expected to build on this basis to boost data sharing and access, and to use even more in all the data spaces.

When asked whether data spaces are relevant to the viability of open data, the panel emphasised that open data is a 'classic example' of the kind of data that plays a role in all data spaces. Open data, including HVDs, is by default available to anyone, including the participants in the common European data spaces, and can be re-used there and anywhere else. Data spaces may also be sources of original open data, whether derived from pre-existing open data (e.g. governments' open data) or created by processing the data space participants' data in ways that make it suitable for re-distribution under an open licence.

The panel stressed the responsibility of the public sector to ensure that open data is used and accessed. Open data is open by definition. Therefore, data users should not need to go through public sector bodies to request the data. The panel concluded that open data is one of the most important resources in the data economy to boost the creation of value ([36]) and that the public sector must ensure they are engaged in all the data spaces and make the data they have openly available.

5. An example of a data space in practice: public procurement data space

PPDS ([37]) is one of the common European data spaces initiatives funded by the digital Europe programme. The purpose of this data space is to connect European databases on public procurement and national procurement datasets available in national portals. Public authorities have huge buying power in the European market. Collectively, they spend EUR 2 trillion every year, or 13.6 % of the European gross domestic product ([37]), on the purchase of services, works and supplies. Some procurement notices by EU bodies are already published through the EU's Tenders Electronic Daily (TED) portal ([38]), run by the Publications Office, to ensure transparency of procedures. However, TED covers only about 20 % of all procurement notices in the EU. The remainder are spread across national or regional levels in different formats, which makes them difficult or impossible to re-use. The intention

---

([35])   https://digital-strategy.ec.europa.eu/en/news/simpl-cloud-edge-federations-and-data-spaces-made-simple.
([36])   https://data.europa.eu/en/publications/open-data-impact.
([37])   https://single-market-economy.ec.europa.eu/single-market/public-procurement/digital-procurement/public-procurement-data-space_en.
([38])   https://ted.europa.eu/.

of PPDS is to create a data space where participating public administrations and the user groups relating to them can offer and consume data describing the procurement processes (Figure 2). It will be voluntary for Member States to participate.
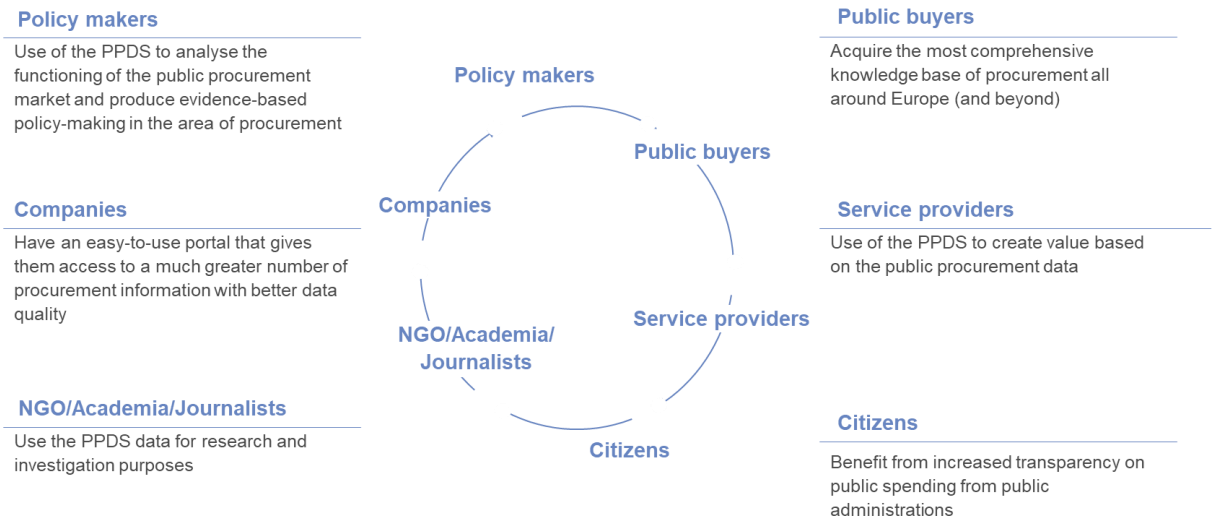
Policy makers
Use of the PPDS to analyse the functioning of the public procurement market and produce evidence-based policy-making in the area of procurement

Public buyers
Acquire the most comprehensive knowledge base of procurement all around Europe (and beyond)

Companies
Have an easy-to-use portal that gives them access to a much greater number of procurement information with better data quality

Service providers
Use of the PPDS to create value based on the public procurement data

NGO/Academia/Journalists
Use the PPDS data for research and investigation purposes

Citizens
Benefit from increased transparency on public spending from public administrations

Figure 2: Envisioned impact on PPDS for various user groups

PPDS has a roadmap to roll out the data space. The starting point for 2023 is to connect PPDS with TED and process basic indicators. The next step for 2024 is to connect interested Member States to PPDS, improve the useability of the portal and include historical data. Beyond 2024, the vision is to connect PPDS to other data sources, both within and outside of the domain of traditional procurement. For example, public procurement touches on health, climate, sustainable procurement and the proper use of EU funds. Public procurement data could also be used to identify collusion. Therefore, the long-term purpose is to enrich the data within the data space and enhance the ways in which the data can be analysed to create insights.

Achieving this vision will require the integration of various data sources. An important factor for integrating various data sources into a data space is having a common format and common understanding of the data. This will allow the data to be shared and made accessible to end users. To help achieve a common understanding of procurement data, the Publications Office and several Member States have developed the eProcurement ontology [39]. The ontology assists in setting common data standards for publishing procurement data through common vocabularies and semantics.

Another success factor for creating a data space is understanding how the data space can meet the needs of participants. To achieve this, the PPDS governance authority holds various co-creation sessions to gather knowledge from various stakeholders such as policymakers, public buyers and small and medium-sized enterprises.

The high-level, conceptual architecture PPDS is designed to provide a single-access point to public procurement data for every user (Figure 3). On the one side, there are data providers, such as TED

---

[39]   https://joinup.ec.europa.eu/collection/eprocurement/solution/eprocurement-ontology.

from the Publications Office and the Member States. On the other side, there are data users with different roles and levels of access to data on PPDS.
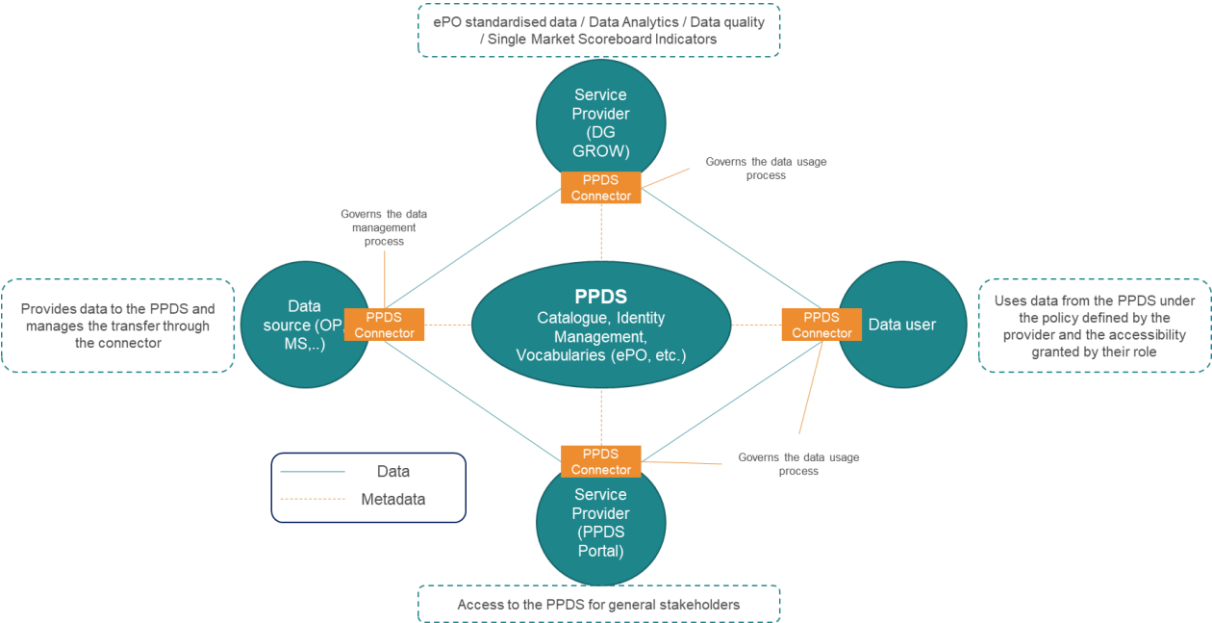

Figure 3: High-level architecture of PPDS to allow interactivity of different types of users who provide or consume data

Open data available on PPDS will be made discoverable on data.europa.eu. Currently, open data from TED is provided as comma-separated values (CSV) files. However, PPDS is based on linked data technology and uses the resource description format (RDF) for its public procurement data. The idea for the future is to provide a SPARQL endpoint ([40]) on data.europa.eu. This will allow users to directly query data that is available as open data on PPDS without leaving the data.europa.eu website. In this sense, data.europa.eu is also a data user of PPDS.

Another prospect for the future is that data.europa.eu users will be able to do combined searches, finding crossovers in data from data.europa.eu with data provided from PPDS. This will require data to be linked in meaningful ways. The panel discussed that PPDS has run a proof of concept like this with wiki data, which enriched the data in the data space. However, the PPDS team also learned that certain data is difficult to link. More work is needed to understand how to link data in meaningful ways. This will become more relevant as multiple common European data spaces are developed. For example, PPDS has so many links to other sectors that it may potentially be linked to many other data spaces in the future. The panel concluded that achieving data linkages would bring added value to the end users in the future, since enriched data enables richer analysis.

In addition to linked data, connecting to other data spaces will require even deeper integrations. In the first stages of creating data spaces, participants within that one data space need to find ways to integrate with each other. The panel hypothesised that this should be facilitated by the fact that the participants typically belong to the same context, industry or use case. However, integrating across data spaces requires a higher level of interoperability that goes beyond the language and ontology specific to one context. Integrations will also be enabled by higher levels of automation. Indeed, interoperability specifics included in the horizontal legislation framework that establish requirements

---

[40]   https://data.europa.eu/data/sparql?locale=en.

for data sharing intend to enable data spaces to 'talk' beyond the boundaries of an individual data space.

Overall, PPDS demonstrates the unique characteristics of data spaces. The panel concluded that PPDS is government-organised but holds large amounts of non-government data. Much of this data is publicly available data, for example the list of organisations and consortia that bid for a contract. However, some information is considered commercially confidential, for example the bidders' rate cards. A data space is a perfect case for a situation where some information may be intentionally published and other information must intentionally be protected. A data space implements the mechanisms by which we can protect what is necessary, open what does not need protection and control who sees what.

6. Support in practice: Data Spaces Support Centre

Data space initiatives require supporting programmes to be successful. One such supporting programme is the DSSC, which is funded under the digital Europe programme. DSSC is operated for the Commission by a consortium of expert organisations. The DSSC contributes to the European strategy for data, specifically by coordinating the development of the common European data spaces. DSSC supports those data spaces that are directly funded by the EU. It fosters (1) the definition of requirements, (2) technologies, (3) processes, (4) standards, and (5) tools to ensure interoperability within and across data spaces.

The goals of DSSC are threefold.

(1) DSSC aims to **create a network of stakeholders.** It does this by supporting the creation of a community of practice in the field of data sharing.
(2) DSSC aims to **identify common requirements** for data spaces. It does this by defining guiding design principles, identifying standards, building blocks, open-source technologies and business models for the creation of data spaces.
(3) DSSC aims to **create a platform for knowledge exchange**. It does this through tools and assets made available for everyone, such as the definition of a blueprint, good practices, templates on model contract clauses, etc.

At the time of writing, DSSC has engaged across the first 11 sectors that the EU is addressing with its funding of data space initiatives. DSSC supports these initiatives by developing various products and services. This includes helping data spaces understand how business models can make them sustainable beyond their EU funding. To develop tools and assets that can help these organisations and initiatives, DSSC holds monthly community meetings and specific workshops with partners to ensure that the users are satisfied with the product and that the product meets their expectations. The panel concluded that these organisations are 'key contributors' to the co-design of the assets being developed. It is important that the different parties feel that they are part of the data space. A second important factor the highlighted by the panel is that data is shared in a trusted manner.

In addition to engaging with the data space initiatives themselves, DSSC engages with other initiatives funded at the EU level that include elements of data sharing. Examples of such initiatives include research and innovation projects funded through Horizon Europe ([41]), cloud-to-edge data initiatives,

---

([41]) https://research-and-innovation.ec.europa.eu/funding/funding-opportunities/funding-programmes-and-open-calls/horizon-europe_en.

initiatives in the area of AI and European digital innovation hubs ([42]). The panel suggested the possibility for DSSC to engage with non-EU funded initiatives in the future, whether created by national governments or the private sector (e.g. the well-known Catena-X ([43]) for the automotive industry). These engagements would be valuable to compare and understand the differences between the data spaces funded by the EU and other similar initiatives.

The panel also emphasised that the 'building blocks' to develop data spaces are not just about technology. Governance, legal and compliance are examples of other types of necessary building blocks. As an example, DSSC is developing an asset called the 'legal compass'. For this asset, DSSC is mapping the different horizontal legislation and policies that establish requirements for the performance of data sharing services. The purpose of this asset will be to help navigate the requirements that data spaces need to comply with.

Finally, through its work on identifying common requirements, DSSC also supports the work of the envisaged EDIB. Specifically, DSSC creates tools to foster and advise on interoperability across data spaces.

7. Data literacy and skills

To make innovations with data possible and ultimately achieve the goals of the European data strategy, the panel concluded that, more than ever, achieving a good level of data literacy in Europe is paramount. Data literacy is the ability of non-specialists to understand the value of data and how it can be used to achieve a purpose. The panel believes that the emergence of data spaces will stimulate a new wave of awareness of the importance of data. The panel predicts that a greater understanding of data by users will also change the way in which data providers offer data. Data providers will be encouraged to supply high data quality. Innovations such as AI technologies may also offer techniques to improve data quality.

To address gaps in digitalisation, the European data strategy includes a chapter on enablers that promote digital skills and data literacy. The Digital Decade policy programme ([44]) also sets targets to ensure enhanced digitalisation across the public sector, citizens and businesses. Furthermore, the Commission has additional funding and policy initiatives to promote digital skills and literacy. For example, Member States can use funding from the Recovery and Resilience Facility ([45]) for digital skills and literacy initiatives. There is also a data space initiative being developed to address the development of skills: DS4Skills ([46]). This data space aims to support the sharing and accessing of education and skills data for purposes such as data-driven policy development, reuse in innovative applications or updates of education and training programmes with emerging skills needs.

The panel concluded that we all need the right tools, skills and knowledge to contribute to the digital transition. This is also important for public sector entities, because they are the ones delivering services to citizens and must be able to deliver these services efficiently in the digital age.

---

[42]   https://digital-strategy.ec.europa.eu/en/activities/edihs.
[43]   https://catena-x.net/en/.
[44]   https://commission.europa.eu/strategy-and-policy/priorities-2019-2024/europe-fit-digital-age/europes-digital-decade-digital-targets-2030_en.
[45]   https://commission.europa.eu/business-economy-euro/economic-recovery/recovery-and-resilience-facility_en.
[46]   https://www.skillsdataspace.eu/.

8. Concepts from data spaces that can inspire data.europa.eu

The panel discussion revealed that data spaces introduce new concepts, practices and models into the world of data sharing, with implications for sharing of open data. The panel highlighted that open data will be a default component of data spaces. As the official portal for European data, data.europa.eu should consider how it can collaborate with the common European data spaces to continue stimulating the supply and reuse of public data resources.

Moreover, data.europa.eu can also draw inspiration from the new ways of working envisioned by data spaces. Even though data.europa.eu is not a data space, the portal does provide a subset of the functionalities offered by data spaces. In particular, the decentralisation, automation, common standards, service focus and user-centricity of data spaces offer concepts that can generate new ideas for data.europa.eu. These are described below.

Each of the following sub-sections is structured in three parts:

(1) a short description of **data space practices;**

(2) **what data.europa.eu could learn and/or adopt from those practices**, assuming that a rationale for the change exists in the interest of data.europa.eu's objectives;

(3) one or more **practical cases** where these practices could be applied to data.europa.eu's current services.

8.1. Decentralisation

The practice of data spaces
Data spaces aim at being decentralised solutions where each participant can offer data using standard technology, under common standards and governance. Owing to standardisation, participants are enabled to move from one data space to another or to participate in multiple data spaces simultaneously with minimal friction. Decentralisation can be considered a sustainable solution for data spaces because it prevents the case where a few participants, whether as providers of data or services, acquire disproportionate power with respect to others. Such a situation could potentially create data monopolies or lock-in situations for the other participants. The upcoming smart middleware for the EU-funded data spaces, Simpl Open, is designed to deal with these situations.

What data.europa.eu could take from this practice
Data.europa.eu offers some services that are similar to the ones typically found in data spaces, such as data catalogues and search functionality. However, data.europa.eu was originally conceived in a centralised manner, as were most data portals. The aim was to be at the centre of the EU's public data resources ecosystem, rather than being just one of multiple services provided by multiple entities. This was a reasonable decision at the time of the original European Data Portal in 2015, as similar services were rare and data.europa.eu was leading the Member States by example. The later consolidation of the EU Open Data Portal (for public data published by EU institutions, agencies and bodies) and the European Data Portal (for the Member States) was also done to have a more standardised and centralised approach when giving access to information.

To apply the data space concept of decentralisation internally, **data.europa.eu could step away from a centralised model** to support the full data sharing ecosystem of public data resources. It is beyond

the scope of this document to fully expand the multiple ways in which this could be achieved, though two practical cases are illustrated below.

A practical case: go beyond the limitations of the traditional national hierarchy for data providers

The hierarchical model that data.europa.eu operates to make data from the Member States discoverable is a legacy of the centralised view. This has been a limitation in the past when, for whatever reason, a local administration in one Member State (e.g. a city) was more active in open data than the central government. The hierarchical model privileged the national data portals, possibly undermining the discoverability of the local administration's data offering.

Over the last year, data.europa.eu has been relaxing this model and aiming at making more data providers discoverable, independently of the hierarchy. The original model should be revised and alternatives evaluated to address its limitations.

A practical case: integrate with third-party service providers for non-mission-critical functionality

data.europa.eu was conceived as a monolithic, centralised system and aims at being self-contained and self-sufficient. data.europa.eu offers all services that the user is expected to need and is not designed to integrate functionality from third parties. Being self-sufficient is reasonable for data.europa.eu's mission-critical functionalities because data.europa.eu is a *public service*. In other words, data.europa.eu's existence prevents, for example, EU citizens from overly depending on commercial search engines (for example, Google) for finding public datasets.

But what about functionalities that are *not* mission-critical? data.europa.eu could offer to integrate third party services into the portal, as long as they respect the terms and conditions set by the Publications Office, which are typically safeguards for the user's privacy and that the code is offered under an open licence.

The dataset preview functionality is an example of where this could be done. The user could either use data.europa.eu's default dataset preview tool or select their favourite tools among the ones published on an 'app store'. Is it then easy to imagine the same being done for data cleansing, data processing, data visualisation and so on.

8.2. Automation

The practice of data spaces

Data spaces also aim at being extremely scalable. Scalability requires processes and procedures to be automated as much as possible and limits the need for human intervention, such as for support services. In turn, automation requires extremely high-quality metadata describing the data being exchanged and all necessary accessory information such as licences, contracts, provenance, lineage and so on.

What data.europa.eu could take from this practice

Automation is similarly instrumental to the sustainability and growth of the model implemented by data.europa.eu to be a pan-European catalogue of public data resources. data.europa.eu has already included a strong element of automation in its process, for example in how metadata is harvested from data providers. Already since 2022, it has been the data.europa.eu team's ambition to be able to grow further by extending its scope to many more data providers than the canonical national providers and EU agencies. To apply the data space concept of automation to data.europa.eu, **data.europa.eu would continue investing in process automation on the portal.**

> **A practical case: improve and be strict about the meta-data quality of data providers as an enabler of automation**
>
> data.europa.eu should aim at achieving higher quality in some of the elements that enable automation. Currently, the quality of the metadata describing the datasets is poor. The metadata quality dashboard (47) shows that one third of datasets are not accompanied by any licence information, and two thirds of datasets do not clearly state a recognisable licence. This would be unacceptable for the operations of a data space. Organisations in a data space that do not produce basic metadata information, such as the licensing of the data they offer, would most likely be excluded from participation.

> **A practical case: automate processes for onboarding more data providers and provide support as you scale**
>
> A practical use case for the evolution of data.europa.eu towards automation is to enable potentially any data provider to make their datasets discoverable on data.europa.eu.
>
> Onboarding many new data providers has several implications. Depending on the success of the initiative, the process of onboarding and supporting new data providers would quickly overwhelm the capacity of data.europa.eu's team. It would also not be financially viable to scale up the 'human' component of that support. Automation, at least for some categories of data providers, would be inevitable. In this model, human support could be maintained for premium providers, such as the EU institutions and the Member States' national portals.
>
> Extending the data provider base would also trigger the need to implement processes that are not performed today. For example, expanding the data providers made discoverable on data.europa.eu brings with it the need to identify and avoid malicious data providers, who may offer datasets of inappropriate or fake data. Only automation can offer viable solutions to these challenges.

8.3. Common standards

The practice of data spaces

Data spaces are open for everyone to provide and use services, if they adhere to the rules by which the data space operates, including the technical standards it adopts. The common European data spaces and the related ambition to become a single market for data aims at using common standards across its component data spaces, ensuring intra- and inter-data space interoperability. The new possibilities that data spaces enable for data sharing also raise requirements for new standards and functionality. For example, trusted data sharing in decentralised systems makes information about the provenance of the datasets (where they come from) and their lineage (what processing the dataset went through) important to determine the integrity of the data. Emerging standards are being developed to provide these functionalities.

What data.europa.eu could take from this practice

Although data.europa.eu is not a data space, the technology and standards it uses belong in the wider space of technology and standards for data publication and application in general. It is a common case for users to use open and non-open data for their reuse cases. It would therefore not be useful that the standards used in the 'open data ecosystem' diverge from the standards used in the 'data spaces ecosystem'.

Indeed, data.europa.eu and the national data portals are becoming more sophisticated, to the point of requiring some of the functionalities more typical of data spaces, for example identity management, access control and high automation, as described earlier in this section. Consider, for example, how

---

(47)   https://data.europa.eu/mqa/?locale=en.

the data portals are preparing to make discoverable also public sector data that cannot be made available as open data, as prescribed by the Data Governance Act.

Furthermore, in the interest of users, it is paramount to guarantee an experience that is as seamless as possible as they move from one context to the other. Causing users to implement different standards to use open and non-open data can only create friction to re-use.

Considering common standards used in data spaces, **data.europa.eu can contribute proactively to the development and promotion of technologies and standards for data sharing,** as it has done for DCAT-AP ([48]).

A practical case: collaborate with initiatives like the DSSC to contribute proactively to the development of technologies and standards for data sharing

data.europa.eu could collaborate with initiatives such as the DSSC, and specifically the processes by which the so-called 'building blocks' are chosen to be recommended to the EU-funded data spaces and other EU data sharing initiatives. We know informally that a few technology standards that are relevant to data.europa.eu, such as DCAT-AP, are already being evaluated. The authors expect many other opportunities for convergence of interests, to whose development data.europa.eu can participate.

A practical case: use your educational platform to showcase standards from data spaces that are also relevant to open data

Since its early days, data.europa.eu has been educating data providers across Europe on best practices and standards for data discovery and publishing. data.europa.eu has a strong audience and is in the unique and privileged position to have pan-European reach. Continuing this legacy, data.europa.eu could volunteer to showcase on the portal the newer and/or less mature standards from data spaces that are also relevant to open data. In this way, data.europa.eu would contribute to the work of DSSC.

For example, C2PA ([49]) is an emerging standard to cryptographically associate metadata to multimedia files in a way that cannot be removed or tampered with. The ambition is to certify, for example, that content was produced by a genuine camera rather than a generative AI engine. This is a countermeasure to the proliferation of deep fakes. data.europa.eu can educate its audience on standards such as this.

8.4. Service-focus and user-centricity

The practice of data spaces
Data spaces will change the behaviour of data re-users in how they search, access and process data. In data spaces, API-like access is the default means of accessing or processing data in the cloud, replacing conventional file downloads of datasets and local processing as the common practice. As part of the decentralised model, re-users will access the data as close to the source as possible, ideally from the data providers themselves rather than from re-distribution venues such as data portals. In this ecosystem, a successful data space is one that is self-sustaining, where users find the services and data they need quickly and reliably.

What data.europa.eu could take from this practice
data.europa.eu is already designed to be service-focused and user-centred. To up the ante in how these concepts of service-focus and user-centricity translate into the data.europa.eu programme, **data.europa.eu should consider what services it is in the unique position to contribute to the data**

([48])   https://joinup.ec.europa.eu/collection/semic-support-centre/solution/dcat-application-profile-data-portals-europe/release/300.
([49])   https://c2pa.org.

sharing ecosystem for public data resources. data.europa.eu has a mandate as an EU-wide public service. This differentiates it from data spaces, most of which are aimed at serving businesses, driven by commercial business models. This grants data.europa.eu unique characteristics. The mission of data.europa.eu is to foster the supply and use of government data. Additionally, only data.europa.eu can be expected to operate a pan-European data search functionality in a way that is transparent and neutral, without being influenced by the commercial interests that characterise commercial search engines.

Accordingly, the way in which data.europa.eu measures its success and decides which services to invest in should be aligned with the purpose summarised in the paragraph above and the changing behaviour of its users. Success factors should centre on whether citizens, research, businesses or other organisations are served in the best way possible to discover data, whether the reuse of open data is stimulated and whether a healthy data ecosystem for public data resources is being enabled. The success of the users in achieving what they need should be the first driver behind service decisions. In other words, the EU is successful not when data.europa.eu is successful per se, but when its citizens, as data re-users, are successfully enabled in general. This change in attitude would enable data.europa.eu to better identify which supportive actions are the most needed and, perhaps, where one cannot expect solutions to be provided by other parties, whether public administrations or private organisations.

A practical case: align your services and measures of success to support the data sharing ecosystem for public data resources

When data re-users are put at the centre, conventional measures of success borrowed from commercial digital business are less relevant (e.g. number of visitors, returning visitors, website stickiness). In this case, other measures of 'ecosystem success' that document how users thrive within the ecosystem become relevant, for example the capability to find the data they need, how fast, how reliably, etc., regardless of the portal or search engine used ([50]). This is particularly true for data.europa.eu, as a public service intended to foster the uptake and supply public data resources (that is, to empower the broader ecosystem).

Data.europa.eu could start experimenting with measuring these elements in how it assesses its success. For example, user satisfaction surveys may not ask the interviewees how easy it is to find good quality e-learning to develop data literacy and skills *on the data.europa.eu portal*, but rather how easy it is to find e-learning *wherever*. Then, if gaps are identified, data.europa.eu would focus its effort on those pain points, e.g. by extending the scope of its academy[51].

A practical case: offer to coordinate with Member States to avoid duplication of effort

Significant effort is invested by the EU and the Member States, at all levels of government, to promote open data in the EU. However, coordination is limited. For example, there is no coordination between data.europa.eu and the Member States to avoid duplication of effort, e.g. in creating new training programmes or developing a necessary piece of open source software needed by one or more portals.

Data.europa.eu may volunteer to facilitate the process by which, in the interest of citizens, the many initiatives are coordinated. The responsibilities of the Public Sector Information Expert Group may be extended to include this objective.

---

[50]   Lego, a Danish toymaker, offers an example of moving from local to ecosystem measures of success, albeit in a different context. Initially, the company measured success by sales store-by-store. In a new model following a digital transition, measures of success included how well the company was selling online across the globe (that is, across the ecosystem), and how effective its software was. data.europa.eu differs from this case by providing a public service and working with an ecosystem beyond its own institution. However, the case study demonstrates the different ways of thinking introduced by ecosystem metrics.

51 https://data.europa.eu/en/academy

9. Conclusions and outlook

Europe has large ambitions and concrete targets for its digital decade to 'empower businesses and people in a human-centred, sustainable and more prosperous digital future' ([44]). A major component of this ambition is a single market for data, along with a legislative framework, common standards and practices that enable data sharing. The panel concluded that data spaces are a priority for the Commission and that data spaces play an important role in achieving the digital transition. Data spaces are also critical instruments to make Europe more resilient and effective.

Open data is an important component of the data sharing ecosystem. In this regard, data.europa.eu has been fostering the uptake and supply of public data resources for almost a decade. However, not all data can be made openly available and non-public data can also benefit the data economy. The data sharing landscape is therefore evolving to enable more complex data transactions.

As a greater number and greater diversity of participants exchange data in new ways, the data sharing ecosystem in Europe changes and the ways by which the ecosystem can be supported evolve. One way to do this is for data.europa.eu to draw inspiration from the concepts used in data space initiatives. The development of solutions for data spaces brings additional practices, technology and legal awareness for how sharing data can take place. These form part of the latest developments and trends in data sharing. Additionally, data.europa.eu should collaborate with data space initiatives to ensure that users of open data continue to be properly served in the evolving data sharing landscape.

The panel also concluded that the success of data spaces will not depend only on technical achievements and the hard skills of processing data. Citizens will need data literacy, meaning that they understand what data sharing can do for their business or public administration office. Going forward, data.europa.eu can take on an educational role to teach its audience about the benefits of participating in data spaces and the missed business opportunities companies could face from a lack of access to data. In this way, data.europa.eu can contribute to the success of data spaces.