CENTRE OF EXCELLENCE
DATA SHARING & CLOUD

Towards Scalable Data Space Interoperability and Federation

A Combined Community-Driven and Participant-Driven Approach, Based on a Common Carrier Layer

Discussion Paper and Call to Action

December 2024

Title:	Towards Scalable Data Space Interoperability and Federation
Authors:	CoE-DSC participants
Editors.	Dr. H.J.M. (Harrie) Bastiaansen (TNO), Dr. G.H. (Gert) Kruithof (TNO)
Date	December 2024
Number of pages:	37 (excl. front and back page)
Number of annexes:	0

Management Summary

The need for data sharing is growing. A major and growing part of that need will be associated with sharing of data across the community borders of individual data sharing initiatives or data spaces. Therefore, the need for an approach for scalable interoperability and federation of data spaces is (rapidly) emerging.

This data sharing use cases which are not confined to specific data sharing communities, have two different origins:

- use cases to support (static or streaming) data sharing across sectoral regional data sharing initiatives or data spaces, and
- use cases to support the new data sharing application areas such as Digital Product Passports (DPPs), Digital Twins (DTs), Privacy Enhancing Technologies (PETs) and event-driven smart contracting approaches.

The investments are considerable to develop mutually interoperable and federated data spaces to support such cross-community use cases. Additionally, they can have substantial lead times, especially as interoperability spans multiple levels: governance, legal, semantic and technical. Moreover, building federation in a bilateral manner is not expected to be scalable when large-scale federation is needed, with many cross-cutting use cases.

Therefore, in this discussion paper we propose an (alternative) approach for large-scale interoperability and federation of data sharing initiatives and data spaces, enabled by recently developed data space standards. The goal is to initiate discussion with the EU Data Space Research and Development communities to jointly work on scalable interoperability and federation of data spaces.

An initial approach is proposed in this discussion paper. It is based on the dynamic establishment of bilateral trusted data sharing relationships between data sharing participants, across the boundaries of specific data sharing communities and data spaces. It enables a new onboarding approach for data service providers and consumers to get mutual access to their data services, with a single point-of-entry and membership, irrespective of the community boundaries of specific data sharing initiatives or data spaces.

This new onboarding model is referred to as the ‘participant-driven’ approach, deployed by means of a ‘Common Carrier Layer’. It builds upon the concept of the ‘Interlinking Layer’ as introduced in the context of the European Mobility Data Space (EMDS), with ‘discoverability’ as key capability to share open data across data spaces or data sharing initiatives. However, as there will be many use cases for which sensitive (non-open) data may need to be shared across data sharing communities, the Common Carrier Layer as introduced in this discussion paper extends beyond the scope of the EMDS Interlinking Layer’ by (1) providing a minimum set of three key capabilities for dynamically establishing data sharing relationships to exchange sensitive data: discoverability, data sovereignty and trust, and (2) providing these three key capabilities ‘horizontally’ to simultaneously support data sharing in and across many other sectors as well.

The Common Carrier Layer for the participant-driven approach is enabled by the new Data Space Protocol (DSP). It allows for bilateral, electronically negotiated and signed, data sharing contracts between participants. The DSP is a game changer as it enables a decentralized implementation for the three above-mentioned capabilities for data sharing: discoverability, data sovereignty and trust.

As elaborated in this paper, the decentralized implementation allows these three capabilities to be provided independent from specific data sharing communities or data spaces. It paves the way for a layered business and technical data space system architecture, distinguishing:

- the participant-driven Common Carrier Layer enabling large-scale federation and interoperability,
- the community-driven data spaces layer aimed at sectoral or application domain specific use cases, and

- the intermediary building blocks and support services layer in which generic capabilities constitute a coherent and complete set of data space support functions, provided as re-usable services, enabling both cost- and time-efficiency in deployment.

The question may arise why the layered business and technical data space system architecture as enabled by the DSP justifies this paper to be positioned as discussion paper and not merely as technical or architectural description. The main reason is that it can have both major legal and architectural impacts, which need to be jointly assessed by the broader data space research and development community on viability, adoption and deployment. These major legal and architectural impacts include:

- The participant-driven approach by means of a Common Carrier Layer covers both *technical interoperability aspects* and *legal bindingness of data sharing transaction contracts*, jointly forming the interoperability and federation fundament for cross-community use cases. It may not be applicable to all types and sensitivity levels of data sharing use cases, but may be adequate for the major part of data sharing needs.
- The participant-driven approach enables individual data sharing participants to become member of the overarching ecosystem of connected participants of the Common Carrier Layer, without being member of a specific data sharing community or data space. Becoming member of the Common Carrier Layer only involves a minimal-threshold, trustworthy, onboarding process based on certified connectors with reliable identification and authentication, under an overarching 'light-weight' trust anchor framework. This onboarding process can be provided as a service by independent Trusted Third Party (TTP) service provider. On its turn, membership to the Common Carrier Layer can be developed as re-usable onboarding service by individual data sharing communities as well, relieving them from this complex task. Vice-versa, an onboarding process for an individual data sharing communities may be re-used for membership to the Common Carrier Layer as well.

Through a combination of both, a two-level membership model (with re-usable onboarding processes) can be supported in which membership of both the participant-driven Common Carrier Layer and community-driven data sharing communities or data spaces coexist.

- The participant-driven approach by means of a Common Carrier Layer *changes the business models and roles*. It provides opportunities for the various roles to focus on their own specific added value by means of separation-of-concerns:
  - For *individual data space participants* to become member (through a low-level onboarding process) of the ecosystem of connected participants of the Common Carrier Layer, with (optionally) lower barriers to join specific data spaces.
  - For *specific community-driven data spaces* to focus on their community specific value adding, without too much effort having to (re-) develop and deploy the generic onboarding and connectivity processes.
  - For *third-party onboarding service providers* to create adequate scale through commoditization, both as Common Carrier Layer for individual data sharing participants that are not bound to (member of) specific data spaces and as re-usable foundation across many data sharing initiatives. Moreover, they may add value in large-scale maintenance, compatibility and interoperability of large numbers of data space connectors.

The approach as presented in this paper may give rise to a fundamental discussion on how it may influence the development and the deployment of the EU Data Strategy and its ambition of the 'Common European Data Spaces'. Therefore, this paper presents recommendations for follow-up, both on governance and business aspects and technical aspects as summarized in the table.

Governance and Business Recommendations	Technical Recommendations
• Create awareness and manage embedding in the development and deployment of the EU Data Strategy. • Do a profound SWOT-analysis on governance and business viability. • Do an assessment of the legal validity across the EU national and EU-overarching laws. • Create alignment of the relation of the DiD / VC architectural approach in view of the (business) guiding principle of a single point-of-entry. • Develop the overarching trust anchor framework for the Common Carrier Layer. • Create awareness and manage embedding in the development and deployment of the EU Data Strategy. • Assess the relevance for, relation with and impact on the emerging data sharing application areas, specifically on DPPs, DTs, and event-driven smart contracting. • Develop a joint collaborative business and governance model for the layered approach. • Develop interconnection strategies and migration roadmaps for existing data sharing initiatives and data spaces to enable large-scale adoption.	• Develop a proof of concept to validate the technical feasibility of the Common Carrier Layer. • Develop an initial set of machine-readable contract templates for electronic negotiation. • Develop adoption and interconnection scenarios and tooling to stimulate adoption of the Common Carrier Layer. • Define interfaces (APIs) for the Common Carrier Layer to simultaneously support many (community-driven) cross data space use cases and application areas.

The technology to enable large-scale interoperability and federation through a participant-driven Common Carrier Layer is becoming available. The main challenge is on its adoption, enabled by an overarching and cross-sectoral business and governance framework. Therefore, we now call upon the EU Research, Development and Innovation communities on federative data sharing and data spaces to jointly proceed working on this challenging and key aspect of the Common European Data Strategy…..

Table of Contents

1 Introduction

The concept of the data space is a corner stone of Europe's Data Strategy [1]. It expresses the ambition of creating the 'Common European Data Spaces' [2] under the supervision of DG Connect. The data space approach provides a scalable mechanism for sovereign, trusted and interoperable exchange of data between many participants and stakeholders.

The data space landscape as it is today consists of both the emerging European sectoral data spaces (as part of the 'Common European Data Spaces' approach) and a multitude of smaller-scale data spaces [3][4].

1.1 Background

Since data sharing use cases and business models are very often not bound to specific communities of participants, it is quite likely that hardly any single formalized data space will cover all the data sharing needs of its participants. Moreover, emerging data sharing application areas such as Digital Product Passports (DPPs) [5], Digital Twins (DTs) [6], Privacy Enhancing Technologies (PETs) [7][8] may need data services provided by parties that are not part of the pre-defined community of participants of specific data spaces.

The adoption of data spaces is driven by both the rapidly growing collection of data sharing use cases and the broadening of emerging data sharing application areas. Combined with the assumption that hardly any single data space will cover all the needs of its participants, a growing need is emerging for establishing dynamic data sharing relationships between participants (data service providers and consumers) in different data sharing communities. Nevertheless, such well-bounded data sharing communities are currently being pursued by many data spaces and (sectoral or application focused) data sharing initiatives: they are designed and organized from the paradigm of 'a data space in isolation'. Therefore, to cater for this need for dynamic data sharing relationships, a paradigm shift is needed towards providing the infrastructure and capabilities for enabling such dynamic establishment of trusted data sharing relationships.

It is expected that data space technology, and specially the recently developed Dataspace Protocol (DSP) [9], will provide the essential legal and technical capabilities to support dynamic establishment of trusted data sharing relationships. As such, the DSP may be the game-changer for the paradigm shift from 'data space in isolation' towards 'dynamic trusted data sharing'. It can provide the foundation for a '*participant-driven'* business model approach to data sharing, allowing for large-scale interoperability and federation through a 'Common Carrier Layer'.

1.2 Scope of the discussion paper

This discussion paper introduces the *participant-driven approach* to data sharing as a scalable approach towards the establishment of dynamic data sharing relationships between participants operating in different data spaces or data sharing initiatives. Moreover, it also supports participation of organizations that are not a member of a formal data space or data sharing community yet.

In its dynamic nature, the participant-driven approach minimizes the need for supporting centralized intermediary services but rather relies on interoperability functions for discoverability, data sovereignty and trust that are provided in a fully-decentralized manner, connecting individual data sharing participants into an ecosystem of data sharing participants, irrespective of the boundaries of specific data sharing communities. As such, this 'bottom-up' *participant-driven* approach is complementary to the more 'top-down' *community-driven approach*, which most often starts from common needs or use cases of a broader community and derives a data sharing solution, including more centralized intermediary building blocks, and derive interoperability and federation approaches thereof to cater. As will be further addressed in this paper, the community-driven data space approach makes it very difficult to support use cases across a multitude of data spaces and data sharing initiatives and, as such, it may

not provide a scalable fundament for interoperability and federation for use cases requiring data sharing between participants across data sharing communities.

The participant-driven approach for large-scale interoperability and federation between data sharing participants is enabled through the concept of a '*Common Carrier Layer*'. It builds and extends upon the concept of the 'Interlinking Layer' as introduced in the context of the European Mobility Data Space (EMDS) [10]. The key capability for the EMDS Interlinking Layer has been defined as 'discoverability', which is key for sharing open data across data spaces or data sharing initiatives. However, it is also to be noted that there will be many use cases for which sensitive (non-open) data may need to be shared across data spaces. To support sharing of such sensitive data, the key capabilities for data sovereignty and trust are needed as well. Therefore, the Common Carrier Layer as introduced in this discussion paper provides the minimum set of three key capabilities for dynamically establishing data sharing relationships to exchange sensitive data: discoverability, data sovereignty and trust. Jointly, these three key capabilities allow a participant-driven business approach to data sharing. They are enabled by the DSP, which can be deployed to guarantee both legal validity / bindingness of data sharing transactions and technical interoperability between data sharing participants across data spaces.

The focus of this discussion paper is on a new, layered, business and technical data space system architecture, enabled by the development of Common Carrier Layer providing the three key capabilities of discoverability, data sovereignty and trust. The new business and technical data space system architecture enables a combined Community-Driven and Participant-driven business model approach, with clear business and operational advantages in evolving towards the ambition of the data economy as defined by the EU Data Strategy.

1.3 Goals of the discussion paper

Based on the ideas of the participant-driven business approach and the Common Carrier Layer as described in the previous section, the goals of this discussion paper are to:

- *Create awareness* of the potential of a participant-driven approach as a complementary approach to the community-driven approach and its importance in support large-scale interoperability and federation between participants across data sharing communities.
- *Explain and elaborate the concepts* of the participant-driven approach and the Common Carrier Layer, their business rationale and legal and technical grounding.
- *Provide input for further discussion and exploration* of the potential business, development and deployment impacts of the participant-driven approach and the Common Carrier Layer for creating more dynamic data sharing relationships as basis for large scale interoperability and federation.

Moreover, this discussion paper does a *call for action* for further elaboration and validation of the business, legal and technical grounding of the participant-driven approach based on the Common Carrier Layer, with a goal for its embedding within the development of both the EU Data Strategy and data space developments.

1.4 Structure of the paper

The following chapter 2 addresses the need for a dual (community-driven and participant-driven) business approach, enabled by the introduction of a Common Carrier Layer and based on the data space connector and dataspace protocols. Chapter 3 elaborates the participant-driven approach with the Common Carrier Layer as enabler for large-scale interoperability and federation of data spaces, together with a base scenario for its deployment. Subsequently, chapter 4 describes the changing business role models with the evolving role of data spaces and service providers in view of the emerging Common Carrier Layer. The overarching conclusions, outlook and recommendations (from both the governance and business perspective and the technical perspective) are provided in chapter 5, after which the final chapter 6 does a call to action upon the European data sharing and data space community to further explore, develop and deploy the new business and technical system architecture as proposed in this discussion paper.

2 An aligned business approach: community-driven and participant-driven federation

The introductory chapter explained that a top-down community-driven approach is unlikely to cover all the data sharing needs of its participating organizations. Neither will it provide a scalable solution to large-scale federation between (participants in multiple) arbitrary data spaces. A bottom-up participant-driven approach, may provide an alternative for sharing data between individual participants, leading to large-scale interoperability and federation between participants across data sharing communities.

The subsequent sections in this chapter will further address the need for a dual (community-driven and participant-driven) approach (section 2.1), the role of the data space connector and dataspace protocol as game changers in enabling the participant-driven approach (section 2.2) by means of the Common Carrier Layer and the potential they provide to develop and deploy a layered business and technical data space system architecture (section 2.3). The final section 2.4 elaborates the need for a joint collaborative business and governance model for the layered business and technical data space system architecture as key for success in realizing the EU ambition of the Common European Data Spaces.

2.1 Need for an aligned business approach: community-driven and participant-driven

The data sharing needs of any participant are in general not limited to the community boundaries of a single data space. This can easily be illustrated by the data sharing needs for both (personal) mobility and logistics. It is clear that both mobility and logistics are by nature cross-border (with a need to share data across geographical areas) and cross-sector (with a need to share data with other sectors as well, such as energy, tourism, smart cities and built environment). However, similar observations of data sharing needs and use cases across the boundaries of specific data sharing communities hold for many (if not all) other data sharing communities as well, either aimed at specific sectoral use cases or application domains. Hence, in view of this broader perspective, the situation that a participant has data sharing needs with use cases that cross the boundary of a single data space community should be seen as the rule rather than the exception.

The data sharing needs and use cases across the boundaries of specific data sharing communities are further increased through the emerging, alternative, data sharing application areas, e.g.:

- The *Digital Product Passports* as mandated by the EC's Ecodesign for Sustainable Products Regulation (ESPR), [5]
- The collaborative *Digital Twin (DT)* [6] in which data from many sectors are used to build an overarching and integrated view and simulation environment,
- The *Privacy Enhancing Technologies (PETs)* [7][8] that are gaining popularity for accessing and processing data in use cases in which the various data sources cannot simply be shared between participant, e.g. due to sensitivity, confidentiality, ethical, privacy or legal issues.
- *Event-driven smart contracting* for data to be shared between organizations by means of a controlled data flow. In logistics, for example, it allows improved visibility along the supply chain and tracking of goods and trucks, and transportation conditions (e.g. for perishable or dangerous goods). Further, it enables the (automated) sharing of transport documents for business reporting or legal compliance. This type of data sharing specifically refers to data sharing concepts and architectures that have been developed by the EU CEF FEDeRATED [11][12].

Although these emerging, alternative, data sharing application areas may at first sight appear to be independent (or even competitive) to data space approaches, it is expected they can benefit from data

space technology as well. Moreover, they have the potential to mutually complement (and even re-inforce) each others strengths. The relative positioning of these emerging, alternative, data sharing application areas has been described in a recent white paper [13]. Although its focus of is on the positioning of PETs and data spaces, the analysis has broader applicability to the other emerging data sharing application areas as well.

Although being well suited for cases where large communities have strong commonalities in their data sharing needs, the 'traditional' and more static and community-driven data sharing application area (with community-specific agreements on legal, governance, semantic and technical interoperability) are less suited for supporting the more dynamic data sharing relationships needs and use cases that cross the community borders. Data space interoperability based on (a multitude) of bilateral agreements between individual data space communities is not considered to be a scalable approach in terms of cross-cutting use cases with participants to be supported, and the number of data sharing communities from which participants might be involved and require federation.

Therefore, an alternative and scalable approach is needed as well. It requires the dynamic establishment of trusted data sharing relationships between participants that are not bound to a single static community and that is scalable across many data sharing communities (data sharing initiatives and data spaces). Moreover, a single point of entry should provide data sharing participants with access to the data services as provided by data service providers across the boundaries of data space communities, i.e. the participant-driven approach.

It is to be noted that the community-driven and participant-driven approach can coexist. Moreover, it will be explained in the remainder of this discussion paper that they can be aligned to mutually reuse each other capabilities, strengthen each other's business proposition, be implemented and deployed through a new (layered) business and technical data space system architecture, and allow individual stakeholders to more effectively focus on its own added value (through separation-of-concerns).

2.2 Game changers: the Dataspace Connector and the Dataspace Protocol standards

The bottom-up participant-driven approach by means of a Common Carrier Layer requires (only) a minimum set of functions to be enabled at the level of the individual participants:

- The ability to communicate the availability of data services with applicable usage conditions by means of standardized (technical) protocols.
- The ability to establish a legally valid / binding data sharing transaction with agreed upon usage conditions prior to sharing sensitive data.
- The ability to verify identities as a trust mechanism for dynamically establishing a trusted data sharing relationship.

In the Common Carrier Layer as basis for the participant-driven approach these functions are provided by means of its key capabilities for dynamically establishing data sharing relationships: discoverability, data sovereignty and trust. They are enabled by the recent development of data sharing standards. As such, these may therefore be referred to as real "game changers".

More specifically, the recently developed data space connector architecture and the Dataspace Protocol (DSP) [9] may provide the (legal and technical) capabilities for the Common Carrier Layer to support the participant-driven business approach [10], enabling dynamic establishment of trusted data sharing relationships across the boundaries of sectoral and application domain data sharing communities or data spaces.

The introduction of a the data space connector architecture together with the DSP are included in the DSSC Blueprint v1.5 [14] and marks a crucial milestone in the development of the European data spaces. They are concisely described in the following paragraphs, subsequently.

2.2.1 The data space connector: separation between control plane and data plane

The data space connector architecture is based on the principle of separation of the control plane and data plane [15], as depicted in Figure 1. It is also referred to as 'out-band' control for federative data sharing.

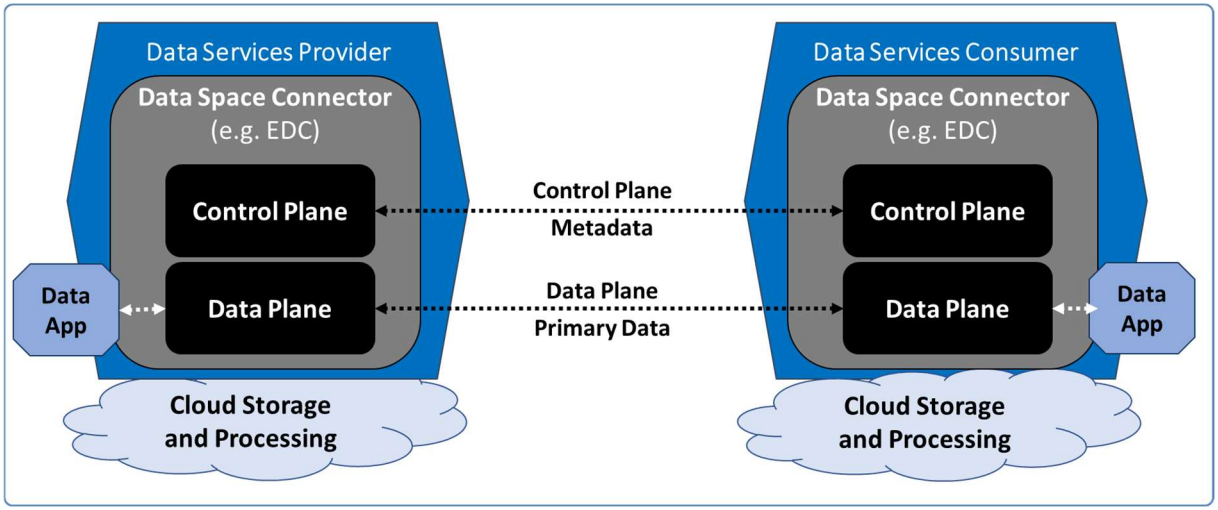

Figure 1: Out-band control for federative data sharing by means of separation of the control plane (for metadata exchange) from the data plane (for primary data transfer).

The *control plane* handles the discovery of the ICT assets offered by connectors and the associated policies. It also handles the contract negotiations. For this, it exchanges metadata with the control plane of other data space connectors.

The *data plane* handles the actual transfer of the shared data with the data plane of other data space connectors. This is referred to as the primary data and can be potentially sensitive data.

A major advantage of using an out-band control mechanism with separation of the control plane and the data plane as depicted in Figure 1 is that it gives flexibility in allowing multiple connectivity protocols at the data plane, e.g. to support multiple types of data sharing and to serve different connectivity needs. Whilst this on the one hand may lead to new interoperability challenges as it allows for differentiation in choices of connectivity protocols to be supported, it is on the other hand expected that only a limited set of connectivity protocols are needed to serve the various types of data sharing, e.g. HTTP, MQTT and Kafka protocols.

Various data space connector implementations are being developed. Currently, the Eclipse Dataspace Components (EDC) [16] attracts major attention for implementing the data space connector according to the separation of the control plane and the data plane as described in this paragraph and for enabling the Dataspace Protocol for interoperability as described in the following paragraph. However, it is to be realized that the EDC is more a software framework for developing data space connectors and less specifying (the architecture and protocols of the) data space connector itself. As such, the EDC leaves open several design choices still to be made on the protocol and data space connector level. This implies that adopting the EDC doesn't automatically imply interoperability with other data spaces that adopt the EDC.

2.2.2 The Dataspace Protocol: interoperability

Sharing data between autonomous entities (participants, data space connectors) requires the provision of metadata to facilitate the transfer of assets by making use of a data transfer protocol. The DataSpace Protocol (DSP) [9] defines how this metadata is provisioned. It is a set of specifications designed to facilitate interoperable data sharing between entities governed by usage control and based on Web technologies. These specifications define the schemas and protocols required for entities to publish data, negotiate usage agreements and access data as part of a federative data sharing architecture or data space, see Figure 2.

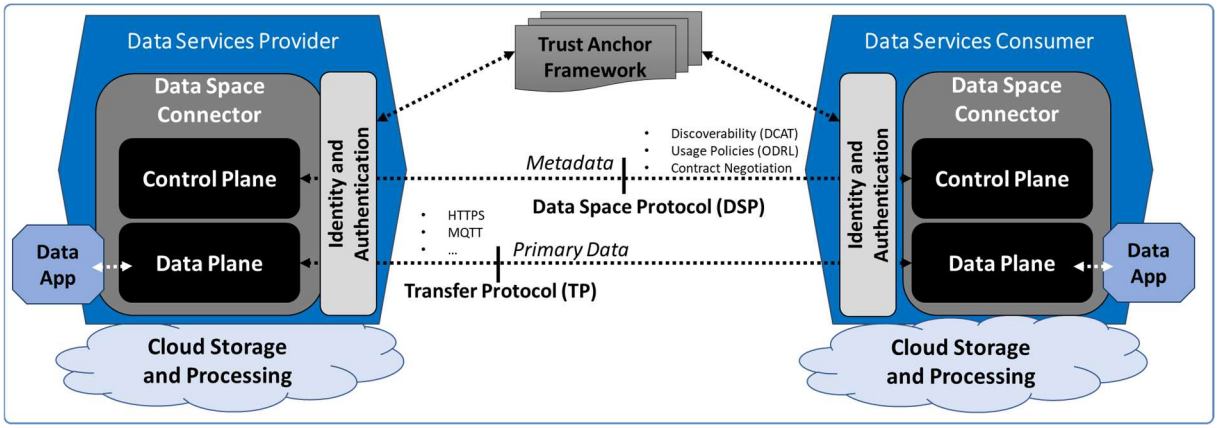

Figure 2: The Data Space Connector and the Data Space Protocol defining the control interface (metadata) between data space connectors.

The Dataspace Protocols define how data assets are deployed (as DCAT Catalogs [17]), how usage control is expressed (as ODRL Policies [18]), how contract agreements that govern data usage are syntactically expressed and electronically negotiated and how data assets are accessed using data transfer protocols.

The Dataspace Protocol specification does not cover the data transfer process itself. While the data transfer is controlled by the Transfer Process Protocol, the data transfer itself (and especially the handling of technical exceptions) is an obligation to the Transport Protocol that will be used.

The fact that the Dataspace Protocol specifications addresses the processes of metadata exchange to enable data sharing, but not the data transfer itself corresponds, to the basic design assumption of separating the control plane (with metadata exchange to enable data sharing) from the data-plane (with the actual transfer of potentially sensitive primary data), as described in the previous paragraph.

Trustworthy identity and authentication management functions are key to establishing a trusted data sharing relationship between the participants in a data sharing transaction. It is to be noted (as depicted in the figure) that these are not within the scope of the DSP. Therefore, an additional framework of (interacting and interoperable) trust anchors providing such trustworthy identity and authentication management services is key to the realization of a large-scale interoperability and federation across data sharing communities and data spaces.

2.3 Towards a layered business and technical data space system architecture

The data space connector architecture and the DSP as described in the previous section allow for decentralization by means of direct peer-to-peer interaction between two data space connectors (participants) of the key capabilities of discoverability, data sovereignty and trust. This minimizes (or even make redundant) the need for centralized services or building blocks to authorize, control and operate legally binding data sharing transactions, which has previously been mainly formalized and operated by data spaces or data sharing initiatives. Hence, this allows for reconsideration of the business approach for developing the EU Data Strategy and its ambition of the Common European data spaces.

Moreover, it enables a new, layered, business and technical data space system enabling large-scale interoperability and federation, combining the potential and benefits of the dual and coexisting community-driven and participant-driven business approach. It is graphically depicted in Figure 3.

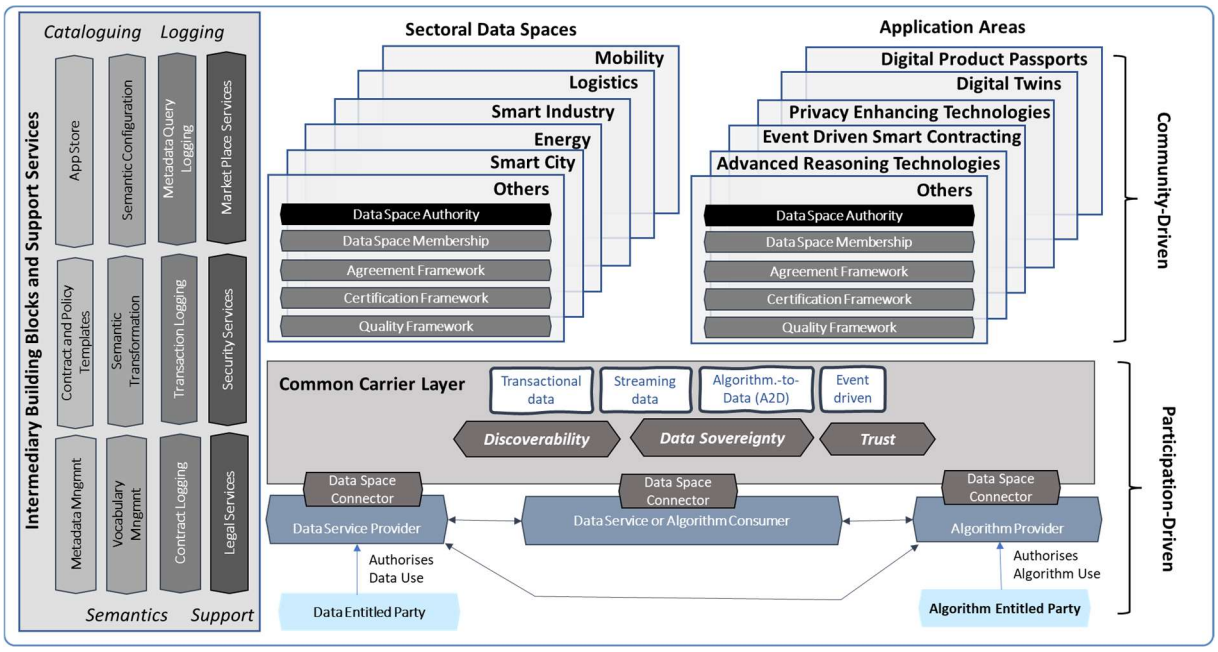

Figure 3: Layered business and technical data space system architecture, enabling large-scale interoperability and federation.

The layered business and technical data space system architecture as depicted in the figure distinguishes three layers: (1) the Common Carrier Layer (participant-driven), (2) the data spaces and application domain layer (community-driven), and (3) the intermediary building blocks and support services layer. These three layers are further addressed in the following paragraphs of this section.

2.3.1 Common Carrier Layer roles: participant-driven

The Common Carrier Layer provides the basic infrastructural capabilities of discoverability, data sovereignty and trust, as depicted in Figure 3. It forms the basis for the participant-driven approach. It underlies (and complements) the ecosystem of community-driven (sectoral or application domain oriented) data sharing communities.

A main goal of the participant-driven Common Carrier Layer is for individual data service providers and consumers to have a single-point-of-access for interconnecting with each other and enabling dynamic data sharing relationships, irrespective of whether they are a member of specific community-driven data sharing initiative or data space.

This participant-driven approach by means of the Common Carrier Layer offers potential opportunities for both individual participants and for data sharing initiatives and data spaces:

- For *individual participants* it provides mutual connectivity and data sharing capabilities in a uniform manner not being bound to closed membership communities of specific data sharing initiatives or data spaces.
- For *individual data sharing initiatives and data spaces* it provides the basic capabilities on discoverability, data sovereignty and trust in a common and reusable manner, relieving them from the to define, develop and deploy these capabilities again for their own specific data sharing initiative. It allows them to focus on their own value added services and capabilities, improves efficiency and effectivity in developing their initiatives through reuse of existing capabilities and prepares them for large-scale interoperability and federation for their participants across the community boundaries of their own initiative.

As also stated in the introductory chapter, it is to be noted that the term ‘Interlinking Layer’ stems from the European Mobility data space (EMDS) initiative. In its communication on the ‘Creation of a common European mobility data space’ in November 2023 [10], the European Commission indicates that the Interlinking Layer is expected to become the core of EMDS. The Interlinking Layer will [citation] ‘*enable the interconnectivity of existing and emerging mobility and transport data spaces and domains. It will notably facilitate the discoverability and accessibility of data from those data spaces and domains*’. As

the EC's communication further states, all the EMDS components will be aligned with the recommendations from the EDIB, the DSSC and relevant building blocks, e.g. provided by Simpl [19].It is obvious that both mobility and logistics are cross-border and cross-sector in nature. Developing a participant-driven approach by means of an Interlinking Layer is therefore of clear interest to the mobility and logistics sector. However, similar observations of hold for many (if not all) other sectors as well, Therefore, the EC's communication has indicated the relevance of developing the interlinking layer to be extendible and interoperable with other sectors as well.

In the participant-driven business approach based on the Common Carrier Layer, the bilateral, legally binding, data sharing contracts can be established between its participants, bypassing or extending (but not conflicting with) agreements made within specific data sharing communities, even when these participants are not member of a specific data sharing community at all or are member of different data sharing communities. As long as identities of counterparts can be adequately authenticated and trusted, any bilateral contracts (e.g. established by means of bilateral contract negotiation as enable by the DSP) will be legally binding.

The participant-driven Common Carrier Layer, its functions and architectural concepts are further elaborated in the following chapter 3.

2.3.2 Data spaces and application domains layer: community-driven

The value proposition a community-driven data sharing initiative (either sectoral or application domain oriented as depicted in Figure 3) is to provide a portfolio of services to support value adding use cases to its participants, enabled by decreasing their investment and operational costs of data sharing. This portfolio is a managed set of services combined with policies, rules, contracts, agreements, and standards. The value adding use cases represent (categorized) data sharing needs of participants.

Within the service portfolio of a community-driven data sharing initiative, a distinction can be made between generic (or use-case agnostic) and use-case specific services. Examples of generic services are, discoverability, identity and trust services. Examples of use case specific services are the provisioning of use-case specific contract templates, data models and domain-specific value adding services.

In a more traditional approach for community-driven data sharing initiatives, both the generic and use-case specific services are designed, developed and deployed grounded in a set of community-specific agreements, i.e. an 'agreement framework'. For generic services however, this may not be the most cost-efficient and business-effective approach.

- On the one hand, reuse of the generic services (potentially operationally offered by third parties) community-driven data sharing initiatives clear *cost- and time-efficiencies* in development and deployment of their initiatives through separation-of-concerns.
- On the other hand, it may provide them a clear *business benefit* in terms of ensured interoperability and federation for the participants beyond the boundaries of their specific data sharing community.

Specifically, building upon generic services for the basic capabilities on discoverability, data sovereignty and trust in a common and reusable manner, allows individual community-driven data sharing initiatives and data space to focus on their own value added services, with improved efficiency and effectivity in development and deployment, and in anticipation of for large-scale interoperability and federation for their participants across the community boundaries.

2.3.3 Intermediary building blocks and support services

As also depicted in Figure 3, Table 1 provides an (non-exhaustive) overview of possible intermediary building blocks and support services. The intermediary building blocks are further categorized into trust building blocks, cataloguing building blocks and semantics building blocks.

Trust	Cataloguing
• *Policy Registry*	• *Metadata Broker*

Manages the policy conditions, e.g. policies for access and usage rights conditions to data services or algorithms as attributed by their entitled parties, including delegation thereof by specific data space participants to other participants. • *Contract Manager* Provides capabilities to conclude, manage and log (legally binding) data sharing transactions, enabling the offering of (potentially sensitive) data services under defined terms and conditions, including for the creation and monitoring of smart data sharing contracts. • *Clearing House* Logs the activities performed in the course of a data sharing transaction. Based on this logging information, the transaction can then be billed. The logging information can also be used to resolve conflicts.	Manages, registers and publishes the ICT-resources available within a data space, e.g. data services, (AI-) algorithms and / or computing resources. • *Contract Template Registry* Provides a set of pre-defined machine-interpretable (ODRL-based) templates for bilateral contract negotiation on (for instance) usage policies and optional components of a legally binding data sharing contract. • *App Store* Manages, registers and publishes data apps that facilitate data processing workflows. These can be deployed within a data space connector.

Semantics	Support Services
• *Vocabulary Hub* Provides facilities for publishing, editing, browsing and maintaining vocabularies (and related documentation), including ontologies, reference data models, schema specifications, mappings and API specifications that can be used to annotate and describe data sets and data services. Vocabulary hubs can be federated to ensuring availability and findability semantic resources. • *Semantic Transformation Engine* Provides semantic transformation services between data formats. It uses vocabularies and mapping specifications, e.g. as provided by a vocabulary hub. The component can be integrated at the data services consumer or data services provider or offered as a service in a data space.	• *Legal services* Offers services to increase the compliance with regulations such as the EU Digital Acts and the GDPR. • *Security services.* Offers monitoring, remote attestation or auditing services in the data space to reduce the risk of cybersecurity attacks. • *Marketplace services* Provides capabilities to create value from data sharing by valorizing data transactions, e.g. through accounting, monetization and / or billing thereof.

Table 1: Intermediary building blocks and support services (non-exhaustive).

2.4 A new business and governance model: collaborative, complementary and coordinated

To fully exploit the potential and benefits of the three-layered business and technical data space system architecture, associated business role and governance models are required. These can be characterized as supporting a collaborative, complementary and coordinated (C3) development and deployment approach, with:

- *Collaboration,* indicating the business approach includes multiple distinctive roles that require alignment through collaboration in jointly developing and safeguarding the individual business interests and viability.
- *Complementarity,* indicating that the distinctive capabilities and roles in the three-layered business approach are mutually consistent, non-overlapping and complete. It provides separation-of-concerns, allowing each of the stakeholders in the three layers to focus on its specific added value.
- *Coordination,* indicating that aligned and agreed upon protocols and interactions are agreed upon across the various roles to enable large-scale interoperability and federation.

The development of such a joint collaborative business and governance model for the layered business and technical data space system architecture may be a key success factor for broad adoption of data space architectures and concepts and realizing the EU ambition of the Common European Data Spaces, as expressed in the EU Data Strategy.

3 Towards a participant-driven Common Carrier Layer

As described in the previous chapters, a participant-driven approach enabling dynamic establishment of trusted data sharing relationships, scalable across many data sharing communities, may serve the needs for cross community use cases and the emerging new data sharing application areas such as DPPs, Digital Twins and PETs.

The Common Carrier Layer as infrastructure enabling the a participant-driven approach may provide a better scalable and common foundation, underlying (and complementing) an ecosystem of community-driven (sectoral or application domain) data spaces.

This chapter further elaborates the participant-driven Common Carrier Layer, its business rationale and capabilities (section 3.1), the approach of legally-binding, bilaterally negotiated data sharing contracts (section 3.2), and a base scenario for how it can provide large-scale interoperability and federation (section 3.3).

3.1 Business rationale for the participant-driven Common Carrier Layer

In describing the business rationale for the Common Carrier Layer as enabler for the participant-driven approach, the paragraphs in this section address the business principles and the main capabilities for the Common Carrier Layer, subsequently.

3.1.1 Business principles for the Common Carrier Layer

The participant-driven business approach and its enabling Common Carrier Layer are driven by the following set of business principles:

1. A participant-driven approach enabling dynamic establishment of trusted data sharing relationships is a main business requirement to support both the needs for cross community data sharing use cases and the new data sharing application areas such as DPPs, Digital Twins and PETs.
2. The participant-driven Common Carrier Layer should enable both for interoperability between participants (data service providers and consumers) active in the same sector and for interoperability with participants active in different / adjacent sectors. E.g., it should allow data space participants to share data across mobility data spaces in different geographical areas and across borders and for interoperability with data spaces serving connected cars, smart cities, tourism, energy and electric charging facilities, logistics, build environment, smart industry, etc..
3. A single point-of-entry and membership provides participants in the Common Carrier Layer access to the data services as provided by other participants, irrespective of the community boundaries of specific data sharing initiatives or data space instances.
4. Dynamic establishment of bilateral trusted data sharing relationships are needed to provide a scalable interoperability and federation approach across the boundaries of specific data sharing communities and initiatives.
5. Alignment is to be taken care of with the concepts, standards and protocols as being defined in the main EU data space reference architecture development and deployment initiatives to spur large-scale adoption and interoperability. Broad acceptation and adoption requires the Common Carrier Layer to be embedded in and aligned with the recommendations from the EDIB, the DSSC, Simpl, and the guidelines of the European Interoperability Framework.

6. The Common Carrier Layer should support multiple types of data sharing. Specifically, the following four types of data sharing should be supported (as described in [20] and has been depicted in Figure 3: (1) persistent (static or semi-static) data, (2) (real-time) streaming data, (3) algorithms for local processing of (sensitive) data, and (4) event-driven smart contracting for data flow control.

Based on these business principles, the main goals of the Common Carrier Layer are to provide a harmonized data sharing infrastructure to enable the participant-driven approach for large-scale interoperability and federation between data space participants (potentially also being member of different data sharing communities.), supporting a broad diversity of data sharing use cases and application areas, both cross-country, cross sector and cross data spaces.

3.1.2 Main capabilities for the Common Carrier Layer

As introduces in section 1.2 and elaborated in the previous chapters, the '*Common Carrier Layer*' is an enabler for the participant-driven approach for large-scale interoperability and federation. It builds and extends upon the concept of the 'Interlinking Layer' as introduced in the context of the European Mobility Data Space (EMDS) [10], which has defined 'discoverability' and 'data access' as the key capability for the EMDS Interlinking Layer, being fundamental for sharing open data across data spaces and data sharing initiatives.

However, as many use cases require sensitive (non-open) data to be shared across data spaces, the key capabilities for data sovereignty and trust are needed as well. Therefore, the Common Carrier Layer as introduced in this discussion paper provides the minimum set of three key capabilities for dynamically establishing data sharing relationships to exchange sensitive data across data spaces and enable large scale interoperability and federation: discoverability, data sovereignty and trust, with:

- *Discoverability,* for the registration, exposure and searchability of both data services (products), participants and connected data sharing initiatives and data spaces
- D*ata sovereignty,* meaning that the entitled parties (e.g. data owners) retain control over their data and can determine who can use it and under what conditions.
- *Trust,* providing assurance mechanisms (at both the legal as well as technical level) to entitled parties (e.g. data owners) for sovereignty over their data.

Jointly, these three key capabilities allow a participant-driven business approach to data sharing. As described in section 2.2 , these key capabilities are enabled by the DSP, which can be deployed to guarantee both legal validity / bindingness of data sharing transactions and technical interoperability between data sharing participants across data spaces.

3.2 Legal embedding: bilateral contract negotiation

As the previous section states, data sovereignty and trust are main capabilities to be supported in the participant-driven Common Carrier Layer. The establishment of legally binding data sharing contracts is a main mechanism underlying the data sovereignty and trust capabilities.

The following paragraphs in this section describe that such legally binding data sharing contracts can be established by the bilateral contract negotiation features of the DSP, with (machine-readable) contract templates included in the bilateral negotiated process, subsequently.

3.2.1 Legally binding bilateral contracts through DSP contract negotiation

The DSSC Blueprint v1.5 [14] identifies two legal building blocks as part of its building block taxonomy [21]: Regulatory Compliance and the Contractual Framework (CF). The CF building block covers contract templates, model clauses and modules that allow transaction participants to regulate and execute specific data transactions.

The DSP as described in section 2.2 provides the option to bilaterally negotiate data sharing contracts. Such bilaterally negotiated data sharing contracts may use and extend the CF Framework of a data

space but can also overrule agreements and clauses formulated in the CF. A simplified illustration on the DSP contract negotiation process steps is depicted in Figure 4.

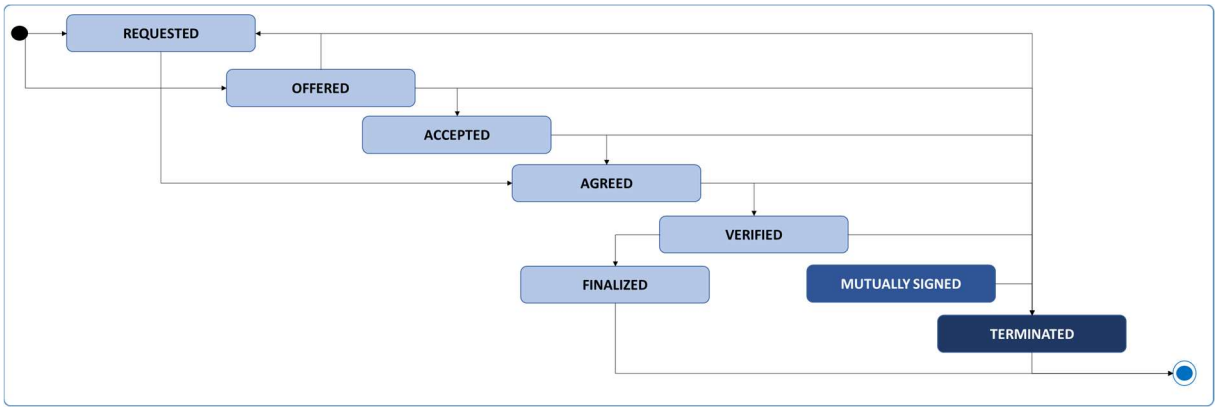

Figure 4: Contract negotiation process steps, simplified version derived from [22].

As the figure shows, an important process step for ensuring legally binding data sharing contracts is the mutual signing by both the data service provider and consumer of the data sharing contract by means of a qualified electronic signature [23]. Providing services for qualified electronic signatures to ensure legally binding data sharing contracts can be an important role of the Common Carrier Layer as part of its trust anchor framework services as depicted in Figure 2. Together with onboarding / membership services, and identification and authentication service, ensuring legally binding data sharing contracts by means of qualified electronic signatures may form a strong and coherent service trust service portfolio.

A studies on the Dutch law [24] (appendix C2), has shown that an electronic, bilaterally negotiated, data sharing contracts is legally binding if it complies to four conditions:

- The electronic contract can be consulted by the contract parties;
- The electronic contract authenticity is assured to a sufficient degree;
- The moment of creation of the electronic contract can be established with sufficient certainty;
- The identity of the contracting parties is established with sufficient certainty.

It is to be noted that although these four conditions for legal bindingness of electronically negotiated data sharing contracts are based on Dutch law in 2020, it was observed that from the European perspective, no additional legislation and guidelines apply (yet). Hence, as next step, an assessment of the validity of these results from the perspective of EU law is required.

3.2.2 Optional components for data transaction contract: contract templates

In addition to the four mandatory conditions for a legally-binding (electronically negotiated) data transaction contract as described in the previous paragraph, there are a number of optional components for a data transaction contract. These are depicted in Figure 5.

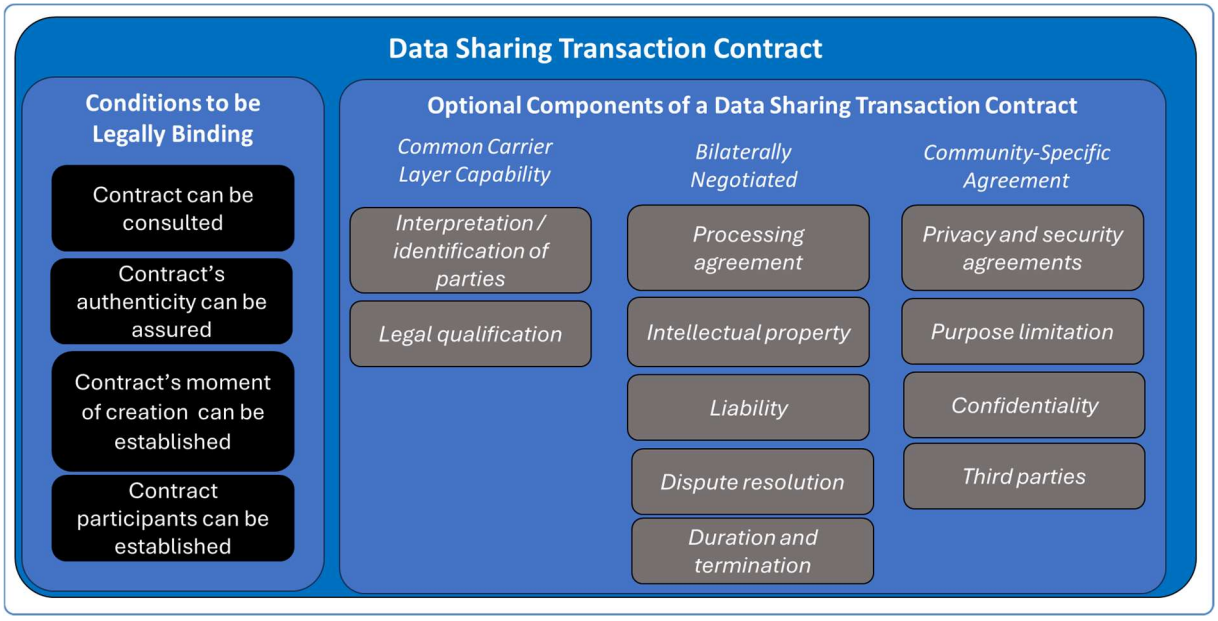

Figure 5: Data transaction contract: conditions to be legally binding and optional components.

As depicted in the figure, the optional data transaction contract components can be categorized according to the layering of the decentralized business and technical data space system architecture as described in section 2.3 and depicted in Figure 3, as either being:

- provided as capability by the Common Carrier Layer,
- agreed upon in a bilaterally negotiated process between participants, or
- addressed as a community-specific agreement within a specific data sharing community.

For the optional data transaction contract components as depicted in Figure 5, Table 2 provides a short description, together with an initial proposition of categorization / positioning as either being a Common Carrier Layer capability, bilaterally negotiated or a community-specific agreement. The initial proposition of categorization in the table may need further discussion.

Common Carrier Layer Capability	Bilaterally Negotiated	Community-Specific Agreement
• *Interpretation / identification of parties* Including aspects such as the name of the parties in the Trade Register, the location details at the Chamber of Commerce, the name and contact details of the (legal) representative. • *Legal qualification* Containing the most common (legal) definitions for the terms used in the agreement, e.g. 'Third Party', 'Agreement', 'Incident', etc..	• *Processing agreement* Stating how another party should handle the (sensitive) data that has been collected. For example, it is recorded how personal data must be secured and for what purposes it may be processed. • *Intellectual property* Describing license agreements and usage rights. • *Liability* Describing the liability for damage suffered by the party that attributably fails to fulfill the obligations under a data sharing transaction. • *Dispute resolution* Identifying the law governing the data transaction contract and the competent judge/court for disputes between parties. • *Duration and termination* Containing the start date and termination date of the data sharing transaction contract, and the procedure for renewal or termination thereof.	• *Confidentiality* Explaining the nature of confidential information and stating the parties' duty of care with regard to the information (even if information is not confidential information). • *Privacy and security agreements* Aiming at taking appropriate technical and organizational measures to ensure a risk-appropriate level of security of the rights of data subjects, including identifying the basis on which personal data may be shared. The most important elements in the agreement are: (i) right to use (sensitive or personal) data, (ii) legality of processing, (iii) accuracy of (sensitive or personal) data to be transferred, (iv) security, (v) compensation, (vi) duration and termination and (vii) liability. • *Purpose limitation* Formulating the objective of collaboration / data sharing. • *Third parties* Describing (the manner in which) third-party services can be obtained, for example with regard to external (legal) advice.

Table 2: Categorization and positioning of the optional data transaction contract components.

For the data transaction contract components that are proposed in the table to be bilaterally negotiated machine-readable, ODRL-based, contract templates can be defined and made broadly available, e.g. by means of Interoperability Profiles managed through a Contract Template Registry (as described in section 2.3). The concept of Interoperability Profiles to support large scale interoperability and federation will be further addressed in the next section.

3.3 The Common Carrier Layer: A base scenario for large-scale interoperability and federation

Based on the development and concepts as described in the previous sections, a base scenario can be made for the development of the participant-driven approach with a Common Carrier Layer that enables large-scale interoperability and federation.

The following paragraphs in this section address the core principles, the core trust capabilities for the Common Carrier Layer and the core Service Provider offering and to enable large-scale interoperability and federation, respectively.

3.3.1 Core principles

The core principles for the Common Carrier Layer to enable large-scale interoperability and federation, are:

1. Participants in the Common Carrier Layer bilaterally control the policies and conditions under which they share their (potentially sensitive) data.
2. The trust and legal capabilities in the Common Carrier Layer are based on bilateral and legally-binding data sharing transaction contracts, which can be dynamically established by means of contract negotiation and ensured to be legally binding by means of qualified electronic signatures.
3. Membership in the Common Carrier Layer is done through a 'light-weight' trust anchor framework which is independent from specific data sharing communities. This trust anchor framework should be broadly accepted. It should therefore be under EC and national member states governance. It may be operated by Trusted Third Party (TTP) providers or an EC-broad mandated organization.
4. A dataspace connector and the DSP are adopted as common standards in the Common Carrier Layer to ensure large-scale interoperability and federation, including:
   - the standardized DSP negotiation process,
   - the ODRL protocol [18] to define usage policies and contracts, including for the optional components of a legally binding data sharing contracts (see section 3.2), and
   - the DCAT protocol [17] to describe and share data services and IT-resources.
5. The development of a limited set of Interoperability Profiles to further simplify the adoption of the Common Carrier Layer and the participant-driven approach. The Interoperability Profiles consist of pre-defined templates for the ODRL-based usage policies and optional components to support legally binding data sharing contracts.

3.3.2 Core trust framework capabilities

As described as core principle for the Common Carrier Layer in the previous paragraph, a 'light-weight' trust anchor framework is required to provide the minimum set of capabilities for trustworthy participant membership of the Common Carrier Layer.

In a base scenario, the three core trust capabilities of a 'light-weight' trust anchor framework in the Common Carrier Layer are the Identity Management capability, the Managed Onboarding capability, and the Negotiated Data Sharing Contract capability. These are subsequently addressed in the following subparagraphs.

The Identity Management capability

For trusted data sharing between two participants, the mutual validation of each other's trustworthiness is important, for which participants need identities that can be authenticated. Identification and authentication of participants must occur at two levels:

- At the level of *legal identities* to identify and authenticate natural persons, organizations or software components as legal entities.

- At the level of *data sharing members*hip, either as member of the participant-driven Common Carrier Layer and/or as member of a community-driven data sharing initiative or data space. It designates legal entities as 'participants', and ensures adherence to the associated onboarding process.

During run-time, a data sharing transaction may include a process for verification of legal identity and status of participants, including their data sharing membership(s).

The identity for participant membership of the Common Carrier Layer reflects that a trustworthy onboarding process has been done, e.g. including connector certification and reliable identification and authentication processes.

Assignment of the participant membership identity (certificate) is part of the overarching trust anchor framework for the Common Carrier Layer, as described in the following subparagraph.

The Managed Onboarding capability

For a participant to become member of the Common Carrier Layer (and, as such, of the participant-driven approach), a basic onboarding process is required to ensure trustworthiness of all stakeholders and assure a basic level of trustworthiness for data sharing transactions. The Managed Onboarding capability should support the issuing and management of membership certificates, which state trustworthiness of participants. On its turn, a membership certificate reflects a successful onboarding process for a participant, e.g. with a check on its legal identity, use of a certified connector, ...

A Trusted Third Party (TTP) can fulfill a service provider role providing membership-as-a-service through a basic onboarding process.

The potential impact of the onboarding process to the Common Carrier Layer on the business (role) models for data spaces will be further addressed in section 0.

The Negotiated Data Sharing Contract capability

As described in section 3.2, the DSP provides the option to bilaterally negotiate data sharing contracts. The bilateral contract negotiation features of the DSP allow (machine-readable, ODRL-based) contract templates to be included in the bilateral negotiated data sharing contracts. An initial set of optional contract templates could be made available by means of Interoperability Profiles, i.e. a core principle as stated in paragraph 3.3.1.

As indicated in section 3.2, a studies on the Dutch law in 2020 [23] (appendix C2) has shown that an electronic, bilaterally negotiated, data sharing contracts electronically signed by means of a qualified signature is legally binding when it adheres to four base conditions. At that time, it was observed that from the European perspective, no additional legislation and guidelines apply (yet) to the data transaction contract. Hence, as next step, an assessment of the validity of these results from the perspective of EU law is required.

3.3.3 Core service provider offering

The elegance of the offering Common Carrier Layer is, of course, its apparent simplicity in enabling the participant-driven approach. It provides a low-cost entry model for both data service providers and consumers, to become members of an overarching ecosystem of data sharing participants, without having to be participants / members of a specific community-driven data space as well.

The core service for the Common Carrier Layer services portfolio is its basic participant onboarding (membership) service, providing access to the data services provided by ecosystem of participants across the borders of individual data sharing communities. This basic membership service can be provided as commodity by Trusted Third Party (TTP) service providers, which are accredited as part of the overarching 'light-weight' trust anchor framework as described in paragraph 3.3.2, indicating that it verifiably adheres to the internationally agreed upon trust-anchor requirements.

As part of the basic membership service, participants are provided with both the verifiable credentials on membership to the Common Carrier Layer and a certified and DSP-compliant data space connector. A DSP-compliant connector may potentially be provided as a managed service including deployment, configuration and maintenance by the TTPs. Moreover, this approach allows the TTP service providers to create adequate scale through commoditization and add value in large-scale maintenance, compatibility and interoperability of large numbers of data space connectors.

The basic membership service for the Common Carrier Layer can be offered by many TTP service providers, e.g. in the ICT and Telecommunications sector. The business model for these service providers could for instance be a combination of a flat data space connector fee with additional service fees for the number of transactions (e.g. on data sets shared, configuration activities, auditing, etc.).

The potential of the Common Carrier Layer and the participant-driven approach in creating changing business role models for data spaces is further elaborated in the following chapter.

4 Changing business role models for data spaces

A participant-driven approach based on a Common Carrier Layer could serve as the foundation for the extended reach of cross-community use cases, enabling large-scale interoperability and federation. This approach addresses both technical interoperability needs and the legal bindingness of data transaction contracts. Although it may not be suitable for all data sharing types and data sensitivity levels, it could effectively meet the data sovereignty, legal and trust needs for many sensitive data sharing use cases.

The introduction of the layered business and technical data space system architecture with a Common Carrier Layer will impact the evolution of current business (role) models for data spaces. Therefore, it may initiate a fundamental discussion on how it can influence the development and the deployment of the EU Data Strategy and its ambition of the 'Common European Data Spaces', giving rise to the following questions that are addressed in this chapter:

- What will be the core and the optional services that a community-driven data space provides to its participants?
- What is the role of a data space in data sharing use cases that span across its community boundaries?

The following sections subsequently address how a participant-driven Common Carrier Layer will enable multiple onboarding models to coexist (section 0), how this can / will lead to changing business role models (section 4.1) and how it may result into variation in community-driven data spaces (section 4.2).

Multiple, coexisting, onboarding models

The Common Carrier Layer has the potential to transform data sharing business (role) models by enabling an alternative, low-level, onboarding process, see also paragraph 3.3.2. It allows individual data sharing participants to join the overarching ecosystem of connected participants within the Common Carrier Layer without needing to be part of a specific data sharing community or data space. Onboarding still requires a trustworthy process involving certified connectors, reliable identification and authentication, under the control of an overarching trust anchor framework. However, this process can be managed by independent Trusted Third Party (TTP) providers rather than specific data sharing communities or data spaces.

Several options can be developed for this onboarding process:

- Third-party onboarding service providers can offer onboarding as a service to individual data sharing communities or data spaces, alleviating them of this complex (and non-core) task.
- Conversely, third-party onboarding service providers can use the onboarding processes of individual data sharing communities or data spaces as a basis for connecting their participants to the Common Carrier Layer.
- A combination of both approaches can support a two-level onboarding process, allowing membership in both the participant-driven Common Carrier Layer and the community-driven data spaces to coexist.

These multiple, coexisting, onboarding models for the community-driven data sharing communities or data spaces and for the participant-driven Common Carrier Layer allow for variations in business roles and data spaces to be developed, as described in the remainder of this chapter.

4.1 Changing business role models with focus on added value

The introduction of the layered business and technical data space system architecture with a participant-driven Common Carrier Layer and the possibility it gives for multiple coexisting onboarding models (as described in the previous section), leads to changing business role models with opportunities for the various roles to focus on their own specific added value through separation-of-concerns:

- *Individual data space participants* can join both the Common Carrier Layer and specific data sharing communities through a single onboarding process.
- *Community-driven data sharing initiatives and data spaces* can concentrate on their community-specific value adding without the burden of developing and deploying generic onboarding and connectivity processes.
- *Trusted Third Party (TTP) onboarding service providers for the Common Carrier Layer* can achieve adequate scale through commoditization, serving as a foundation for (and across many) data sharing initiatives and for individual data sharing participants. They can add value through large-scale maintenance, compatibility, and interoperability of numerous data space connectors.

These roles with focus on their own specific added value allow for a variation in community-driven data space approaches to emerge.

4.2 Variation in community-driven data space approaches

As the previous section describes. the layered business and technical data space system architecture (see section 2.3) allows roles to focus on their own specific added value, which, on its turn, enables a variation in community-driven data space approaches. Two scenarios at both end of the spectrum can be distinguished, see Figure 6: the Minimum Viable Data Space (MVDS) approach and the Full Service Data Space (FSDS) approach.

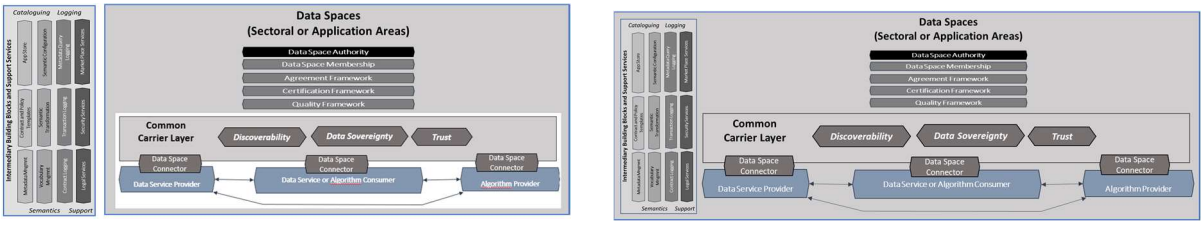

Figure 6: The Minimum Viable Data Space (MVDS) approach (l) and the Full Service Data Space (FSDS) approach (r).

In the MVDS approach, as depicted in the left side of Figure 6, the capabilities are provided by the independent service providers according to the layered data space system architecture as described in section 2.3. Specifically, onboarding processes with the key capabilities for discoverability, data sovereignty and trust are provided by a service provider for the Common Carrier Layer. The role of a community-driven data space is minimized to only providing data space specific and value adding core functions under control of a data space authority. The data space does not provide intermediary building blocks or support services, such as the cataloguing, semantics and logging services as depicted in the figure. For these, participants can individually choose to use services that are provided by independent service providers, external to the various community data spaces.

In the FSDS approach, as depicted in the right side of Figure 6, the data space provides the complete set of capabilities to its participants, i.e. for each of the layers as described in section 2.3. The data space has a Governance Authority that has defined a set of policies and rules in the data space, is responsible for the on- and offboarding of participants and provides intermediary building blocks and support services.

As indicated, the MVDS and the FSDS represent approaches at both ends of the spectrum. A multitude of intermediate or mixed approach may emerge as well, e.g. in which either the onboarding process capabilities and / or the intermediary building blocks and support services are partly provided by the community-driven data spaces and partly by Trusted Third Party (TTP) service providers.

5 Conclusions, outlook and recommendations

Throughout this discussion paper, many potential benefits have been addressed for introducing a participant-driven approach and its enabling Common Carrier Layer to support large-scale interoperability and federation in the emerging and overarching ecosystem of data sharing initiatives and data spaces. As such, it can become the fundament underlying the EU Data Strategy and its ambition of the 'Common European Data Spaces'.

This concluding chapter of the discussion paper provides the overarching conclusions (section 5.1), an outlook towards a multi-service data space system architecture (section 5.2) and the recommendations for follow-up, both from the governance and business perspective and the technical perspective (section 5.3).

5.1 Overarching conclusions

The need for a scalable approach to interoperability and federation between data spaces is driven by the growing need for use cases across the community borders of specific data sharing instances or data spaces. This growing need is based on both use cases to support data sharing across sectoral regional data sharing initiatives or data spaces and on use cases to support emerging new data sharing application areas: DPPs, DTs, PETs and event-driven smart contracting approaches.

In our experience, the discussions in the development and deployment of specific 'Common European Data Spaces' build upon the (implicit) assumption that data spaces are developed to be community-driven, taking sector specific use cases as basis. As such, they each develop a 'customized' approach, addressing the broad set of capabilities across each of the three layers of the business and technical data space system architecture (as described in section 2.3) in an integrated and potentially 'siloed' manner, i.e. the Common Carrier Layer, the data spaces layer, and the intermediary building blocks and support services layer. This results in the Full Service Data Space (FSDS) approach, see section 4.2 and Figure 6. This, on its turn, implies the need for a federation approach across each layer as well, being only feasible on a bi-/multilateral basis between individual data spaces, by means of an architectural approach such as 'federable intermediary building blocks, data space proxies or gateways, etc.'. The more data sharing communities or data spaces are involved, the more complex and costly federation on such bi-/multilateral basis becomes. Therefore, this is not considered as a scalable solution for large-scale interoperability and federation.

This discussion paper therefore proposes an alternative approach for large-scale interoperability and federation. It is based on the dynamic establishment of bilateral trusted data sharing relationships between data service providers and data service consumers, enabled through the DSP-standard, and deployed by means of a Common Carrier Layer. It is expected to provide a (more) scalable interoperability and federation approach across the boundaries of specific data sharing communities and data spaces. Moreover, it should enable a new onboarding and membership model for individual participants, based on a single point-of-entry that gives data service providers and data service consumers mutual access to each other's data services, irrespective of the community boundaries of specific data sharing initiatives or data spaces. This new onboarding and membership model is referred to as 'participant-driven'.

It is to be noted that interoperability on intermediary building blocks and support services is a separate topic and should not be mixed up with the basic interoperability and federation approach as presented in this paper. For instance, in a use case where two participants in two different data spaces want to share data, there is no need for semantic interoperability between the data spaces but between the two participants, for which a third-party Vocabulary Hub service which is independent from any specific data sharing initiative or data space may be adequate.

5.2 Outlook: towards a multi-service data space system architecture

The numerous cross-community use cases and emerging data sharing application areas (as described in section 2.1) lead to a growing demand for cross-community data sharing. They provide the motivation for the participant-driven business approach based on the Common Carrier Layer as introduced in this discussion paper.

These cross-community use cases and emerging data sharing application areas may differ in their business and operational requirements, the manner in which they are governed and the interaction patterns they need in their data sharing processes. Nevertheless it is to be noted that they also have shared basic data sharing commonalities: they may require sensitive data to be shared between participants as result of processing or reasoning tasks. This results in a common need to support controlled and trusted sharing of sensitive data, which defines the basis for harmonization by means of the Common Carrier Layer. This is illustrated in Figure 7.

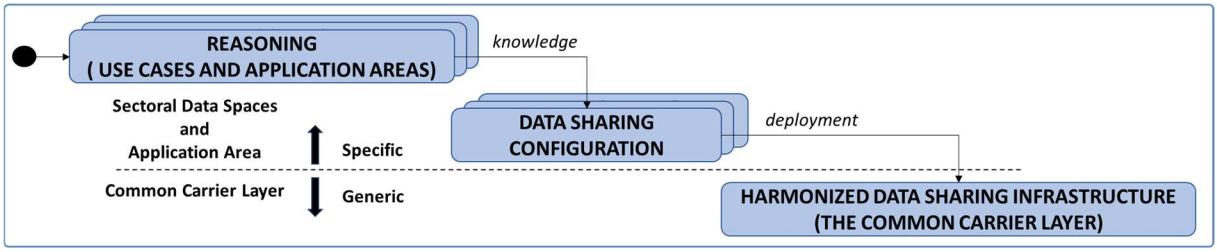

Figure 7: Common process activities from specific use cases application area needs towards generic deployment on a harmonized multi-service data sharing infrastructure: the Common Carrier Layer.

As the figure shows, the role of the generic Common Carrier Layer is to provide a harmonized and reusable foundation simultaneously serving the data sharing needs of multiple, cross-community, data sharing use cases and application areas. It supports the separation between a generic infrastructure layer and a multitude of application areas (or services), requiring a well-defined interface to expose the functionalities of the generic Common Carrier Layer to be accessible, configurable and manageable to the cross-community data sharing use cases and application areas.

It is to be noted that such an approach for “opening generic infrastructure” to be accessible, configurable and manageable by external services is not new. Most notably, similar concepts of opening the generic infrastructure have previously been developed and adopted by the telecommunication sector for opening their infrastructures. They have defined the OneAPI interfaces [25] (previously referred to as “Parlay-X” interfaces) as standard for communications service providers (CSPs) to expose their (fixed or mobile) networks for third-party configuration, jointly creating an open service environment [26].

A similar approach can be pursued by means of the Common Carrier Layer with a set of well-defined interfaces to simultaneously support multiple cross-community data sharing use cases and application areas over a harmonized infrastructure. The roadmap towards such a multi-service data space system architecture based on the Common Carrier Layer needs to be further developed, with specific focus on both the governance models addressing the associated new business role models and stakeholder interests, the enabling interfaces to simultaneously deploy the variety of cross-community data sharing use cases and application areas and the alignment of the operations processes across the business roles, specifically the on- and offboarding processes and the data sovereignty and trust processes (e.g. on usage and access control).

5.3 Recommendations

The potential for a participant-driven based on a Common Carrier Layer to enable large-scale interoperability and federation may be evident. Nevertheless, further assessment and development on the business, legal and technical viability are necessary. Therefore, the following paragraphs provide recommendations for next steps on both the governance and business aspects and on technical aspects, respectively.

5.3.1 Governance and business recommendations

The governance and business recommendations on the further assessment and development of a participant-driven Common Carrier Layer include:

- *Create awareness and manage embedding in the development and deployment of the EU Data Strategy.*

  The Common Carrier Layer as introduced in this discussion paper provides the key capabilities of discoverability, data sovereignty and trust to enable large-scale interoperability and federation across data spaces. Broader awareness on the relevancy and the concept of the participant-driven Common Carrier Layer across the landscape of sectoral data sharing initiatives is needed. Moreover, its concepts may need embedding within the EU Data Strategy with its ambition of the Common European data space. This holds for both the development of its reference architectures and for the actual deployment in the Common European Data Spaces. A community of forefront stakeholders with a common goal to strive for adoption together with EC representatives on the overarching EU Data Strategy may take the lead.

- *Do a profound SWOT-analysis on governance and business viability.*

  Based on the ideas as provided in this discussion paper, a more profound assessment on the governance and business viability of the participant-driven Common Carrier Layer approach is necessary. This may be done by means of an in-depth SWOT (Strengths, Weaknesses, Opportunities, and Threats) analysis. The SWOT analysis should address, amongst others, the scope and limitations of the participant-driven Common Carrier Layer approach and the pro's and con's in comparison to alternative approaches for large-scale federation and adoption.

- *Do an assessment of the legal validity across the EU national and EU-overarching laws.*

  As stated in section 3.2, an assessment on the legal-bindingness of electronically negotiated and signed data transaction contracts (as foundation for the participant-driven Common Carrier Layer approach) has previously been done in 2020 on the basis of the Dutch law. A similar assessment on their legal-bindingness is therefore needed across the overarching EU and its member states laws and jurisdictions.

- *Create alignment of the relation of the DiD / VC architectural approach in view of the (business) guiding principle of a single point-of-entry.*

  In paragraph 3.1.1, a single point-of-entry is stated as business principle for participants to access the data services and interact with other stakeholders in the federation of data spaces as enabled by the Common Carrier Layer. At the same time, the EU promotes the use of Distributed Identities (DiDs) by means of Verifiable Credentials (VCs) as a recommended standard to implement claims and their attestations. On this aspect the business principle of a single point-of-entry and the technical proposal of the use of VCs may be or become contradictory. This potential contradiction / conflict needs to be further addressed as part of the alignment of business and technical architecture development for large scale interoperability and federation.

- *Develop the overarching trust anchor framework for the Common Carrier Layer.*

  As addressed in paragraph 3.3.2, a 'light-weight' trust anchor framework is required to provide the minimum set of capabilities for trustworthy participant membership of the Common Carrier Layer: the Identity Management capability, the Managed Onboarding capability, and the Negotiated Data Sharing Contract capability. The role of a highly decentralized architecture based on the DSP and extended with a. Identity Management capability should be developed for the Common Carrier Layer. For Identity Management, technologies for Self-Sovereign Identities (SSI), Verifiable Credentials (VCs) and Distributed IDentities (DIDs) should be considered. Existing trust frameworks should be assessed on their suitability for providing an overarching trust anchor framework for the Common Carrier Layer, e.g. those as enumerated in [20] (section 8.4): the Gaia-X trust framework [27], the DSBA trust framework [28] and the iSHARE trust framework [29].

- *Assess the relevance for, relation with and impact on the emerging data sharing application areas, specifically on DPPs, DTs, and event-driven smart contracting.*

  Emerging new Data Sharing application areas such Digital Product Passports (DPPs), Digital Twins (DTs), Privacy Enhancing Technologies (PETs), and event-driven smart contracting approaches

are not confined to specific sectoral or regional communities. As described in section 2.1, these emerging data sharing application areas can benefit from data space concepts, protocols and technology. Moreover, they have the potential to mutually complement (and even re-enforce) each other's strengths. These mutual relevance and relationshops should further be assessed.

- *Develop a joint collaborative business and governance model for the layered approach.*

  The layered approach and the introduction of a participant-driven approach with a Common Carrier Layer allow for reconsideration of the business approach for developing the EU Data Strategy and its ambition of the Common European data spaces. Business roles are changing and new mutual business and technical interdependencies may arise. Therefore, a new business approach may be needed, which has been referred to in section 2.4 as being collaborative, complementary and coordinated (C3).

- *Develop interconnection strategies and migration roadmaps for existing data sharing initiatives and data spaces to enable large-scale adoption.*

  Many data sharing initiatives already exist for which it may be of interest to join in a participant-driven Common Carrier Layer approach. Such existing data sharing initiatives have their own implementations, for which it may not be (business- or technology-wise) feasible to update and migrate them towards a layered data space system architecture based on a Common Carrier Layer as described in this paper. Nevertheless, low-impact evolution and migration approaches (e.g. based on gradually introducing the dual onboarding options for participants in their data sharing initiative) may provide an attractive option for interconnection. Such, interconnection strategies and associated migration roadmaps for existing data sharing initiatives and data spaces may be key for large-scale adoption of the participant-driven Common Carrier Layer approach.

5.3.2 Technical recommendations

The technical recommendations on the further assessment and development of a participant-driven Common Carrier Layer include:

- *Develop a proof of concept to validate the technical feasibility of the Common Carrier Layer.*

  To assess, demonstrate and disseminate the concept of the Common Carrier Layer and the technical viability of its underlying architecture, capabilities and building blocks, a proof-of-concept needs to be developed. The proof-of-concept should include a (limited number) of illustrative and representative use cases and contract negotiation aspects with associated Interoperability Profiles with contract templates.

- *Develop an initial set of machine-readable contract templates for electronic negotiation.*

  As described in section 2.3, machine-readable (ODRL-based) contract templates can be defined and made broadly available through a Contract Template Registry for the data transaction contract components that can be bilaterally negotiated. This can be done for instance for the conditions as proposed in Table 2 for the processing agreement conditions, the intellectual property conditions, the liability conditions, and the conditions for duration and termination of contracts. Further identification and assessment of a set of relevant contract templates needs to be developed and made commonly available.

- *Develop adoption and interconnection scenarios and tooling to stimulate adoption of the Common Carrier Layer.*

  The Common Carrier Layer will need to be developed, deployed and embedded in the EU Data Strategy and data sharing landscape. To stimulate adoption of the Common Carrier Layer by existing and emerging data sharing initiatives and data spaces across sectors and application domains, the barriers for interconnection should be made as low as possible. This can be enabled by defining both representative scenarios for adoption and interconnection and developing supporting tooling.

- *Define interfaces (APIs) for the Common Carrier Layer to simultaneously support many (community-driven) cross data space use cases and application areas.*

  Interoperability between the various layers enabled by the emerging Common Carrier Layer should be enabled through generic Application Programming Interfaces (APIs) for large scale adoption.

This allows data sharing instances and data spaces to access the capabilities provided by the Common Carrier Layer in an easy manner. Well-defined APIs are needed for the Common Carrier Layer, enabling communication patterns to simultaneously support (community-driven) cross data space use cases and application areas. Such a similar, “multi-service”, layering approach has previously been developed and adopted by the telecom industry to expose their (fixed or mobile) networks for third-party configuration, jointly creating an open service environment.

6 Call to action

The enabling technologies, standards and building blocks for realizing the EU Data Strategy and its ambition of the ‘Common European Data Spaces’ are rapidly maturing. The introduction of a Common Carrier Layer has been proposed in this discussion paper as basis for a participant-driven approach to enable large-scale interoperability and federation within the emerging ecosystem of data sharing initiatives and data spaces, addressing both business, legal and technical aspects. Although there are still challenges to address, none of them seems fundamentally intractable or a fundamental research topic.

Data spaces are deployed now and operated by businesses, handling sensitive data in an appropriate way, but (to our opinion) too much developed in a siloed, and therefore non-interoperable, manner. The real challenges on large-scale interoperability and federation therefore evolve from the technical towards the business and governance realm, with focus needed on collaborative business models in which development, deployment and large-scale adoption of the participant-driven approach and the Common Carrier Layer are secured.

We now call upon the EU Research, Development and Innovation communities on federative data sharing and data spaces to jointly proceed working on this challenging and key aspect of the Common European Data Strategy…..

References

[1] European Commission (2020). 'A European Strategy for Data'. https://digital-strategy.ec.europa.eu/en/policies/strategy-data.

[2] European Commission. 'Common European Data Spaces'. https://digital-strategy.ec.europa.eu/en/policies/data-spaces.

[3] International data spaces Association (IDSA). 'Data Spaces Radar'. https://internationaldataspaces.org/adopt/data-spaces-radar.

[4] EMDS CSA PrepDSpace4Mobility (2023). 'Inventory of Existing Data Ecosystems'. https://mobilitydataspace-csa.eu/wp-content/uploads/2024/03/2024-03-19-deliverable-d2.1-inventory-of-data-ecosystems-v3.pdf.

[5] European Commission (2022). 'Ecodesign for Sustainable Products Regulation'. https://commission.europa.eu/energy-climate-change-environment/standards-tools-and-labels/products-labelling-rules-and-requirements/sustainable-products/ecodesign-sustainable-products-regulation_en.

[6] TNO. 'Reliable and Real-Time Digital Twin Systems'. https://www.tno.nl/en/digital/digital-innovations/data-sharing/reliable-real-time-digital-twin-systems.

[7] OECD (2023), 'Emerging Privacy Enhancing Technologies – Current Regulatory and Policy Approaches', OECD Digital Economy Papers, No. 351. https://www.oecd.org/publications/emerging-privacy-enhancing-technologies-bf121be4-en.htm.

[8] The Royal Society (2023), 'From Privacy to Partnership – The Role of Privacy Enhancing Technologies in Data Governance and Collaborative Analysis'. https://royalsociety.org/topics-policy/projects/privacy-enhancing-technologies.

[9] International Data Spaces Association (IDSA). 'Dataspace Protocol - 2024-1'. https://docs.internationaldataspaces.org/ids-knowledgebase/v/dataspace-protocol.

[10] European Commission (November 2023). 'EC Communication - Creation of a Common European Mobility Data Space'. https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:52023DC0751.

[11] EU FEDeRATED project. 'EU-Project for Digital Cooperation'. https://www.federatedplatforms.eu.

[12] EU FEDeRATED project (2024). 'FEDeRATED Reference Architecture'. https://www.federatedplatforms.eu/images/Library/Activity2/FEDeRATED_Reference_Architecture.pdf.

[13] Dutch Centre of Excellence for Data Sharing and Cloud (CoE-DSC), Big Data Value Association. 'Leveraging the Benefits of Combining data spaces and Privacy Enhancing Technologies - Joint BDVA and CoE-DSC White Paper'. https://coe-dsc.nl/wp-content/uploads/2024/04/CoE-DSC-BDVA-Whitepaper-Leveraging-the-Benefits-of-Combining-Data-Spaces-and-Privacy-Enhancing-Technologies.pdf.

[14] EU Data Spaces Support Centre (DSSC) project (2024). 'The DSSC Blueprint v1.5'. https://dssc.eu/space/bv15e/766061169/Data+Spaces+Blueprint+v1.5+-+Home.

[15] EU data spaces Support Centre (DSSC) project. 'DSSC Blueprint v1.5- Control plane versus Data plane'. https://dssc.eu/space/bv15e/766067960/Data+Exchange.

[16] Eclipse Foundation. 'Eclipse Dataspace Components'. https://projects.eclipse.org/projects/technology.edc.

[17] World Wide Web Consortium (W3C) (2020), 'Recommendation - Data Catalog Vocabulary (DCAT) - Version 2'. https://www.w3.org/TR/vocab-dcat-2.

[18] World Wide Web Consortium (W3C) (2018), 'Recommendation - ODRL Information Model 2.2'. https://www.w3.org/TR/odrl-model.

[19] European Commission (EC). 'Simpl: Cloud-to-Edge Federations Empowering EU Data Spaces'. https://digital-strategy.ec.europa.eu/en/policies/simpl.

[20] European Mobility data space Coordination and Support Action PrepDSpace4Mobility (EMDS CSA PrepDSpace4Mobility, September 2023). 'Deliverable D3.1 - Towards a Common European Mobility Data Space - Perspectives, Recommendations and Building Blocks'. https://mobilitydataspace-csa.eu/wp-content/uploads/2024/03/2024-03-19-deliverable-d3.1-analysis-report-v3.pdf.

[21] EU Data Spaces Support Centre (DSSC), "DSSC Blueprint v1.5 - Taxonomy of building blocks". https://dssc.eu/space/BBE/178421909/Organisational+and+Business+building+blocks.

[22] International Data Spaces Association (IDSA). 'Dataspace Protocol - Specification'. https://docs.internationaldataspaces.org/ids-knowledgebase/v/dataspace-protocol/contract-negotiation/contract.negotiation.protocol.

[23] Wikipedia. "Qualified electronic signature". https://en.wikipedia.org/wiki/Qualified_electronic_signature.

[24] NL AIC Data Sharing Working Group. 'Responsible Data Sharing in AI' 2020. https://nlaic.com/wp-content/uploads/2020/10/Responsible-data-sharing-in-AI.pdf. The part on the assessment of the legal validity of electronically negotiated data sharing contracts is only contained as appendix C.2 in the (extended) Dutch version of this report: 'Verantwoord datadelen voor AI'. https://nlaic.com/wp-content/uploads/2022/02/Verantwoord-datadelen-voor-AI.pdf. *In the Dutch Language.*

[25] Wikipedia. "OneAPI (GSM telecom)". https://en.wikipedia.org/wiki/OneAPI_(GSM_telecom).

[26] Raivio, Yrjo & Luukkainen, Sakari & Seppälä, S. (2011). Towards Open Telco - Business Models of API Management Providers. Proceedings of the Annual Hawaii International Conference on System Sciences. 1 - 11. 10.1109/HICSS.2011.454.

[27] Gaia-X European Association for Data and Cloud (2022), "Gaia-X Trust Framework", https://gaia-x.gitlab.io/policy-rulescommittee/trust-framework.

[28] Data Spaces Business Alliance (DSBA). 'Unleashing the European Data Economy'. https://data-spaces-business-alliance.eu.

[29] iSHARE Foundation. 'iSHARE – Trust Framework for Data Spaces'. https://ishare.eu.

www.coe-dsc.nl
info@coe-dsc.nl

Colofon

This discussion paper is a result of the work of the Dutch Centre of Excellence for Data Sharing and Cloud (CoE-DSC, www.coe-dsc.nl), specifically its work on interoperability and federation of data spaces.

The goal of this report is to sparkle the discussions and initiate a call-to-action for the complex challenges of large-scale interoperability and federation of data spaces and data sharing initiatives, a key success factor for broad adoption of data space architectures and concepts and realizing the EU ambition of the Common European Data Spaces, as expressed in the EU Data Strategy.

Contributors

Participants in the CoE-DSC community

Editors (TNO):

- dr. H.J.M. (Harrie) Bastiaansen
- dr. G.H. (Gert) Kruithof

Contact

The Centre of Excellence for Data Sharing and Cloud (CoE-DSC)

E-mail — info@coe-dsc.nl

Website — www.coe-dsc.nl

December 2024.