

DS4Skills: Towards a Data Space for Skills Blueprint

This report provides a high-level overview of research activities and key results in identifying key needs of the stakeholders and their ongoing and preferred solutions on key topics and the first feedback for the Data Spaces Support Centre (DSSC).

This document summarises key conclusions on the necessary actions and system changes that can directly inform the development of the Skills Data Spaces Blueprint.

It builds on top of the Data Space for Skills (DS4Skills) Work Package (WP) 3 – “Developing the blueprint for the set-up and the future deployment of the data space for skills” - where the preferred future in relation to the skills data space is explored by analysing the learnings from WP2 – “Preparing an inventory of existing platforms” and conducting additional research through in-depth interviews and analysis (T3.1 – Identifying key needs of the stakeholders and their ongoing and preferred solutions on all key topics).

To that end, several group interviews and co-creation sessions have been conducted with key stakeholders from 20 relevant skills data initiatives to understand their design approaches and needs in relation to the use of and access to skills data. Through these interviews and co-creation sessions, key use cases and requirements for the innovative applications of the preferred future were also envisioned.

The service design approach that was adopted provides a systems-thinking and user-centred methodology to understand initiatives from the following dimensions:

WHAT: The solutions, services and applications that these initiatives are developing as part of skills data spaces.

WHY: The motivations and drivers for the creation of skills data spaces.

HOW: How initiatives are building skills data spaces, including critical areas and challenges.

1. Research Results

The analysis of our research reveals several challenges and opportunities to strengthen the development of a data space for skills, and highlights important learnings across seven key themes:

Theme 1. Human-centred approach

A human-centred approach to sharing skills data is seen as a fundamental principle and a motivating force for the initiatives interviewed for this report. It is viewed as a core mechanism for creating value in a skills data space and reflections on this prompted the articulation of five aggregate use cases, whereby a skills data space adds value by enabling, mapping, matching, and forecasting skills and competencies.

Theme 2. Semantic interoperability

Interviewees emphasised the need for semantic interoperability, enhanced data quality, and the possibility of producing AI systems that could operate with incomplete data. All these aspects together could define a Minimum Interoperability Mechanism covering both functional and technical issues.

Theme 3. Essential technical elements

There was broad agreement among interviewees regarding the set of building blocks needed to enable the skills data space. In addition to the standard building blocks for data spaces, this analysis identified an additional need for 'human centricity' and AI as specific building blocks in skills data spaces. In addition, there was a clear need articulated to identify Personal Data Intermediaries as defined in the European Data Governance Act and in the data spaces architecture (DSSC, GAIA-X, IDSA).

Theme 4. Business models

Most initiatives interviewed for this research are at an early exploratory stage of business model development, and the majority of them are funded by public subsidies. While there was broad agreement that data space services should be free to individuals, there was diversity in whether interviewees emphasised profit or contributions to a 'greater common good' in their business models. This is linked to initiatives' different positions and relationships with regard to public sector institutions such as universities.

Theme 5. Complexity of data ecosystems

The novelty of the data ecosystems approach is quite evident, as there are no standardised complete governance models yet, only initial general framings. A majority of the initiatives expressed a preference for a decentralised data ecosystem, but there is a lack of clear reference models. The level of complexity in these data ecosystems is highly dependent on the number of actors involved, and this is why many initiatives are starting quite narrow in scope. The absence of proper governance rules and agreements for data spaces management was identified as another fundamental and crucial challenge at present.

Theme 6. Readiness and growth strategy

The main challenges and vulnerabilities hindering the rollout and growth phase are: financing, governance, implementation of legal frameworks, participation and engagement by the different ecosystem actors, lack of or reduced political support, lack of common vocabularies, lack of a strong selling point, and the complexity of agreements and contracts among the different ecosystem actors.

Theme 7. Innovative services and applications

Analysis of the different services and applications provided by the initiatives supported the consolidation of a high-level value proposition for individuals interacting with a skills data space, and distinguished according to four distinct categories of services.

2. Deep-dive into the essential technical elements

The aim of this activity was to explore the building blocks used by the interviewed initiatives, determine which ones can currently be used, and define the ideal future building blocks that the project can use to achieve its objectives. In this chapter, key concepts that were identified for the different building blocks are provided.

2.1 Interoperability building blocks

In the interoperability building blocks category, interviewees often mentioned difficulties such as specificity of technical language and duplication of services. They recommended the use of standardised vocabularies for semantics, data models, and APIs to improve interoperability. Specific examples of standards currently in use include JSON-LD, Open Badges v2 (OBv2), and Schema.org for data representation.

For ideal future building blocks, interviewees suggested the implementation of a universal plugin that would integrate different solutions, the use of the W3C Verifiable Credentials standard, and a mechanism for digital wallet interoperability. They also recommended the use of JSON-LD as the appropriate way to describe skills data models and pivot skills ontologies that support broad semantic interoperability among the knowledge domain. Instead of imposing a single ontology, the interviewees agreed on the development of semantic translators to and from pivot ontologies as the right approach. Other recommended future standards to improve interoperability are GraphQL API and Open Badges v3 (OBv3).

2.2 Trust building blocks

Security, anonymity, pseudonymity, explainability, consent, and contracts were identified as key factors for the trust building block. Interviewees recommended the

use of secure data architectures and the management of consent and contracts based on standards such as Kantara Consent Receipt and Open Digital Rights Language (ODRL) to make processes and outcomes easily explainable. Standards such as IDP, SSO, and Verifiable Credentials were identified as the main standards in use today.

For ideal future building blocks, interviewees recommended the use of decentralised protocols, self-sovereign identity (SSI) management, smart contracts, and decentralised AI training. The interviews also identified the need for building blocks that could guarantee the individual's control over their data through Personal Data Intermediaries.

2.3 Data value building blocks

Data accessibility, comparable data, transparency, and interoperability of metadata and data were all highlighted as key elements of the data value building block. Interviewees recommended collaboration on a common vocabulary to standardise specifications for describing datasets and services in the data space. The Gaia-X catalogues of services, organisations, and datasets, as well as insight reports as JSON dashboards, were mentioned as the main standards used today.

With respect to the ideal future building blocks, interviewees recommended the use of non-fungible tokens (NFTs) for gamification, FAIR principles to optimise the reuse of data, and data vault modelling.

2.4 AI building blocks

Interviewees expressed the need for building blocks to evaluate algorithms against reference ethical criteria. Their recommendation was to follow internationally recognised ethical guidelines for AI, such as those provided by the Organisation for Economic Co-operation and Development (OECD) or the United Nations Educational, Scientific and Cultural Organization (UNESCO). As ideal building blocks for the future, interviewees suggested the use of AI modelling platforms, ethical AI, edge AI translators, and image recognition support.

2.5 Governance building blocks

Interviewees agreed on the need for a multi-level governance system, ranging from general agreements to specific data sharing contracts. The Sitra Rulebook was recommended as a model for general agreements and data sharing contracts. Interviewees also suggested a public/private governance model and policies that encourage cooperation and reciprocity in the ecosystem to avoid vendor lock-in. The role of data intermediaries and personal data intermediaries (PDI) was emphasised, as well as the need to mutualise efforts to develop common building blocks, as the non-profit Prometheus-X has done.

In terms of business building blocks, interviewees stressed the need for a new building block to track the value contributed by each participant in the data space in order to redistribute the value generated accordingly.

3. Synthesis of conditions for success

This section summarises the key needs that the initiatives expressed during the interviews process. The aim is to create a concise list of needs that can be directly used to create the Blueprint for the Skills Data Space and for other DSSC activities. Interview data included perspectives about what actions and changes would be necessary to enable the development of a data space for skills in line with the principles described above and in the EU data strategy.

Here those perspectives are consolidated into **nine conditions for success**, each of which are briefly described, and inform development of the skills data space Blueprint as described in the following sub-section.



Figure 1: The nine conditions for success

Permission to experiment: Data spaces represent a new idea that is both complex and ambiguous. Moving from this idea to operational solutions is an uncertain process that requires experimentation. Some such experiments, when successful, have granted confidence to the larger endeavour, but more experiments are needed to bring the practice of data spaces to maturity. This process can be accelerated by facilitating a culture in which initiatives, such as those interviewed here, are encouraged to iterate and experiment, and in which failure is accepted and learned from.

Cultivate champions: Interviewees consistently highlighted the importance of commitment and confidence among the individuals working to design and develop data spaces. The individuals in key organisations that are driven by a personal passion to build the foundations of data spaces were described as diamonds to be found, cherished and encouraged. Such champions come from all disciplines but share a common vision. They are committed to spreading the word, ensuring funding, building experiments and solutions. Without such champions, data spaces will not happen.

Unity of the small: One of the interviewees observed that data spaces are for ‘second stringers’. Big players that dominate the market are often not interested in sharing data, either because they don't need to or because they believe that sharing data may limit their competitiveness. Instead, data spaces can empower medium and smaller players by increasing their access to data and enabling the creation and delivery of better services for their users and customers. Collaboration and data sharing provide smaller actors with a stronger position in this emerging field.

Broad awareness and recognition: The complexity and novelty of data spaces makes a proper definition and value proposition challenging. Broad awareness and acceptance of the basic value proposition is nonetheless essential to ensure uptake and use of data spaces and related services as they emerge. This, in turn, will require that the idea of a data space is clearly communicated to stakeholders, enablers, decision-makers and policy-makers, and this will require investments and effort.

Renewal: Moving from limited and closed data models to decentralised and open data ecosystems requires massive transformation. Legacy systems can inhibit such transformation at technical, legal, operational and functional levels, and the deep changes required to implement data spaces are often not welcomed. The need for renewal is particularly pronounced in regard to legal frameworks, in part due to the truism that legislation lags behind technological development, but also because of the significant number of legal regimes and jurisdictions with which a data space will interact. Effective development of a data space will require renewal of legislative frameworks, public sector agencies and public-private collaborations.

Public-private partnerships: Public funding has been essential to preliminary work so far conducted to develop data spaces, but there is a concern about financial dependence on the public sector. While interviewees acknowledged the necessity of public funding to create public infrastructures and sandboxes, there was a clear desire to see developments towards the self-sustainability of initiatives, in order to secure the overall health of the emerging data spaces. Future developments of data spaces are likely to be driven by public and private partnerships, which will require

negotiations and agreements regarding funding, revenue and profit sharing in consortia.

Collaboration on infrastructures and building blocks: Data spaces will only function to the extent that there are sufficient and functional mechanisms to facilitate the transmission and sharing of data across between participants in the space. It is well recognized this requires infrastructure and common building blocks, but work in this area to date is episodic, siloed, and marked by redundancies. A public-private EU data infrastructure consortium would be one mechanism to facilitate collaboration across relevant countries and companies to share knowledge, tools and processes in order to optimise resources, reduce costs and improve the efficiency of early engagement in data spaces.

Interoperability: Interoperability is fundamental for a data space to operate. While significant attention has been paid to technical interoperability in the early development of data spaces, there is a clear need to build the foundations of legal, semantic, functional, and operational interoperability in order to enable actors with different systems, legacies and set-ups to work together. It is also important to consider interoperability at the level of the ecosystem, and not only in-between actors in the ecosystem.

Trustworthy system: It is critical to create a reliable, safe and well-governed space. At this stage, trust is most relevant when considering the interactions between actors in the data ecosystem, but as data spaces mature it will be critical to establish systemic trust. This will inevitably be derived from how data spaces are governed and how that governance is communicated. For instance, it is critical that only validated actors are admitted into the data space, as one malevolent or unreliable actor can undermine trust in the entire system.

4. Recommendations for the DSSC

Noting the Data Spaces Support Centre (DSSC)'s important role as an authoritative reference point for developments around European common data spaces, its capacity to coordinate and mobilise activities and knowledge sharing between engaged stakeholders, and its connections with European regulatory bodies, DS4Skills recommends that the DSSC take the following actions:

1. **Clearly articulate and promote a concrete value proposition** of data spaces in general, in specific sectors and to various stakeholder groups. For skills data spaces, that value proposition should reflect and highlight the complementarity and variety of value offerings summarised in Figure 2 and Figure 3. Clearly communicating this type of value proposition will be important for mobilising and engaging the support and engagement of key stakeholders as in the *implementation of the forthcoming blueprint*. The DSSC may wish to consider conducting or contracting additional research to assess to what degree these value propositions resonate across other domain and sector-specific data spaces.

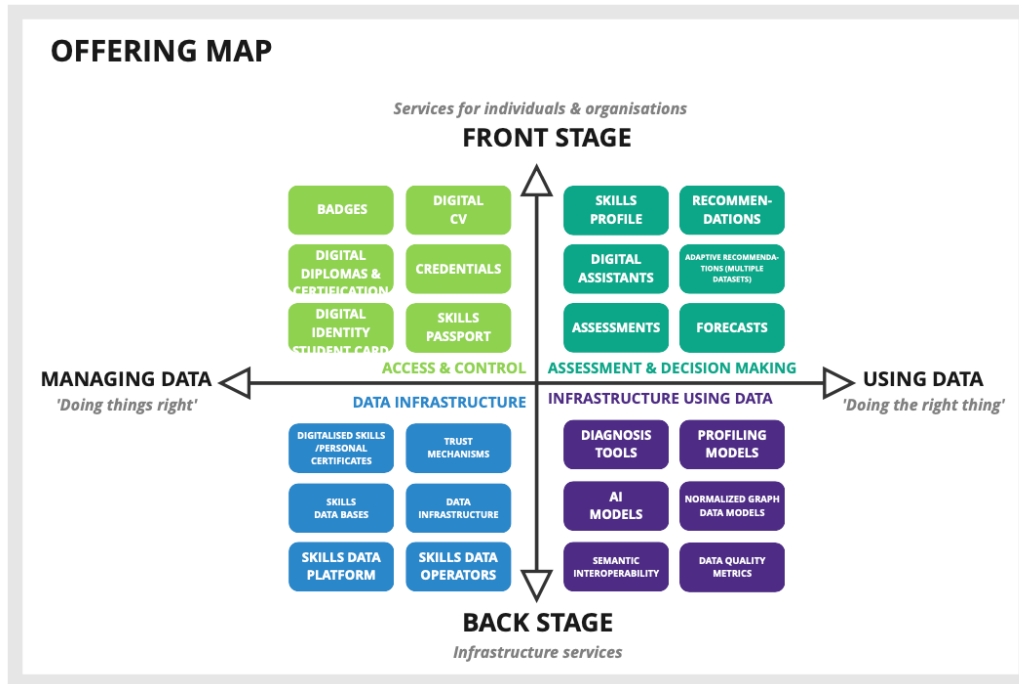


Figure 2: Offering map of innovative services and applications

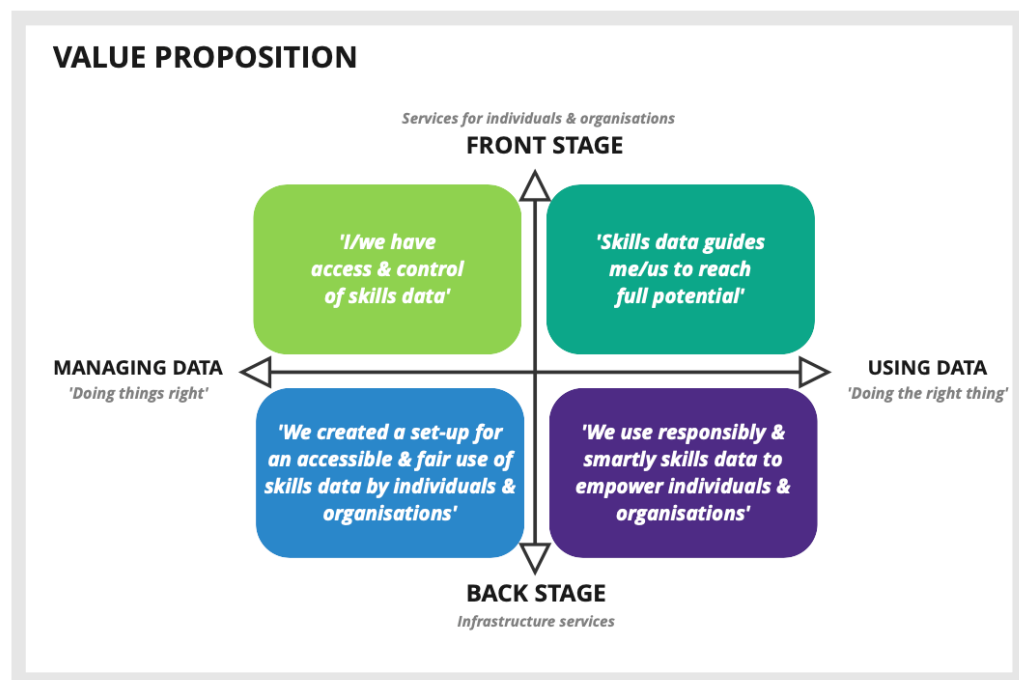


Figure 3: The four cornerstones of the value proposition of Skills Data Spaces

- 2. Clearly articulate and promote the importance of a human-centric approach** to data spaces in its core communications and resources. As noted in this report's analysis, the initiatives included in this research view a human-centric approach as a critical technical and governance element of emerging ecosystem architectures. Furthermore, it's also important for building trust and creating value for the people and organisations that skills data spaces serve. DSSC is encouraged to clearly communicate that a human-centric approach is essential for all data spaces with which individuals or personal data will interact. As such, it should be explicitly integrated already in the early design phases. The DSSC may also wish to encourage the EDIB to instruct its technical subgroup to prioritise a cross-cutting human-centric approach. To ensure this, the research established the need for specific tools and building blocks to enable Personal Data Intermediaries ensuring people have control over their data and organisations comply with those requests. Today the DSSC work only focuses on non personal data sharing which is sub optimal as most if not all data spaces will require personal data sharing.
- 3. Collect and share success stories and learnings of what *not* to do**, especially those relevant to early builders and entrants into data spaces. This should emphasise stories that can inspire and mobilise new and already engaged stakeholders, but it should also include valorising examples of challenges and failures in order to foster a culture of experimentation and co-learning that is necessary to advance data spaces concept and practice towards maturity. DSSC may wish to conduct or contract research to collect such stories and is encouraged to disseminate them widely.
- 4. Continue to support the coordination of small-and-medium commercial actors** to strengthen their market position and efficiency in a market that is currently dominated by large technology actors with outsized resources and influence. The interviews conducted by DS4Skills highlight the value of DSSC's coordination, and the potential it has to maximise coordination and collective influence, what we defined as 'unity of the small'.
- 5. Facilitate public-private budgeting and strategizing exercises.** Next to the more appealing dialogue about the business rationale and foreseen benefits for data spaces, the DSSC should actively support structured dialogue between key public and private actors in order to assess the long-term costs and financial implications of maintaining data spaces. Sustainability and cost-mitigation strategies should be developed and informed by suitable business models and growth strategies that DS4Skills is prepared to discuss with DSSC.
- 6. Encourage the European Data Innovation Board (EDIB) to address legal interoperability** and especially confidence in consistent application across EU Member States. Once formed, the EDIB will play a key role in advancing cooperative European governance of data sharing and data spaces. The issue of legal interoperability is so complex and entrenched that the EDIB should be encouraged to consider creating a standing body, such as a roundtable of member states. This body should foster an ongoing dialogue and co-learning between competent authorities, in order to assess and advance progress on

legal interoperability for data sharing and data spaces. The DSSC may also wish to encourage the EDIB to instruct its subgroup of competent authorities to prioritise these issues in its work.

- 7. Encourage the European Commission to support experimentation** in developing data spaces and the services they will support. DSSC may wish to consider recommending specific mechanisms, such as regulatory sandboxes or the use of synthetic data, specifically to establish environments in which private and public sector actors can experiment in order to enhance value offerings, without fear of non-compliance.

- 8. Define where some technical building blocks should fit in the proposed architecture.** During the interviews, that focused also on the technical building blocks based on the OpenDEI architecture, some functionalities emerged that were difficult to fit in the existing proposal. We would encourage the DSSC to propose how User Experience (UX) as key component for human-centricity, data analytics and artificial intelligence building block would fit in the overall data spaces blueprint.

www.skillsdataspace.eu



**Funded by
the European Union**

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Commission. Neither the European Union nor the granting authority can be held responsible for them.