

Data Spaces between Co-ordination Challenges, Financial Viability and Public Policy



Data Spaces between Co-ordination Challenges, Financial Viability and Public Policy

This work was approved and declassified by the Digital Policy Committee (DPC) on 10 June 2026.

This document, as well as any data and map included herein, are without prejudice to the status of or sovereignty over any territory, to the delimitation of international frontiers and boundaries and to the name of any territory, city or area.

Please cite this publication as:

OECD (2026), *Data Spaces between Co-ordination Challenges, Financial Viability and Public Policy*, OECD Publishing, Paris, <https://doi.org/10.1787/0066f11b-en>.

ISBN 978-92-64-76919-9 (print)
ISBN 978-92-64-95733-6 (PDF)
ISBN 978-92-64-36362-5 (HTML)

Photo credits: Cover © Prostock-studio/Shutterstock.com.

Corrigenda to OECD publications may be found at: <https://www.oecd.org/en/publications/support/corrigenda.html>.

© OECD 2026



Attribution 4.0 International (CC BY 4.0)

This work is made available under the Creative Commons Attribution 4.0 International licence. By using this work, you accept to be bound by the terms of this licence (<https://creativecommons.org/licenses/by/4.0/>).

Attribution – you must cite the work.

Translations – you must cite the original work, identify changes to the original and add the following text: *In the event of any discrepancy between the original work and the translation, only the text of the original work should be considered valid.*

Adaptations – you must cite the original work and add the following text: *This is an adaptation of an original work by the OECD. The opinions expressed and arguments employed in this adaptation should not be reported as representing the official views of the OECD or of its Member countries.*

Third-party material – the licence does not apply to third-party material in the work. If using such material, you are responsible for obtaining permission from the third party and for any claims of infringement.

You must not use the OECD logo, visual identity or cover image without express permission or suggest the OECD endorses your use of the work.

Any dispute arising under this licence shall be settled by arbitration in accordance with the Permanent Court of Arbitration (PCA) Arbitration Rules 2012. The seat of arbitration shall be Paris (France). The number of arbitrators shall be one.

Foreword

Data spaces have attracted substantial public and private investment as a means of enabling trusted data sharing among multiple parties, yet few have reached operational and financial viability. This report examines the novel use of data spaces where alternative data-sharing mechanisms are insufficient. To that end, it draws on evidence from 235 initiatives listed on the International Data Spaces Association (IDSA) Radar (as of February 2026). In addition, it examines case studies of successful, on-track and failed data spaces, and highlights an OECD Expert Workshop held at the Data Spaces Symposium 2025 in Warsaw (Poland) on 11 March 2025.

The report focuses on three main areas. It identifies four boundary conditions that define the class of use cases for which data spaces hold a comparative advantage. It also analyses the bootstrapping challenge that helps explain much of the maturity and adoption gap. Finally, it sets out eight policy considerations for closing the gap between public investment, and the operational and financial viability of data spaces.

Through its analysis, the report clarifies where data spaces are structurally necessary and how they can become financially self-sustaining. In so doing, it contributes to the OECD's ongoing work on data flows, governance and privacy, including on data free flow with trust (DFFT). At the same time, it supports implementation of the OECD Recommendation on Enhancing Access to and Sharing of Data [[OECD/LEGAL/0463](#)].

This report benefitted from the input of the Working Party on Data Governance and Privacy of the Digital Policy Committee (DPC), as well as from the OECD DFFT Experts Community, including through the March 2025 OECD Expert Workshop and written comments on the draft. It was drafted by Christian Reimsbach-Kounatze, Senior Economist and Policy Analyst at the Data Policy, Privacy and Online Safety (DPPS) Division. The report also benefitted from the contributions of Shunya Muromachi, Policy Analyst, as well as feedback from Maiko Meguro, Senior Policy Analyst, under the supervision of Clarisse Girot (Head of the DPPS Division). The research and stakeholder engagement provided by Shihori Maeda, former Policy Analyst, is gratefully acknowledged. This work was made possible by the generous contributions of Japan.

Table of contents

Foreword	3
Abbreviations and acronyms	6
Executive summary	7
1 Overview	10
1.1. Objective: Assessing the distinctive value of data spaces	11
1.2. Scope and approach	13
References	15
Endnotes	17
2 Understanding data spaces	18
2.1. From data integration to data governance	19
2.2. Emerging definitions and converging architectures	20
2.3. A two-tier working definition	21
2.4. Data spaces and related concepts	24
References	26
Endnotes	28
3 Multi-party co-ordination as the distinctive value proposition and challenge of data spaces	29
3.1. The co-ordination problem behind multi-party data sharing	30
3.2. Comparative advantage	32
3.3. How data spaces address the co-ordination problem of multi-party data sharing	35
3.4. The bootstrapping problem as central barrier to viability and adoption	40
References	46
Endnotes	49
4 The evolving policy landscape	53
4.1. Regulatory frameworks that require multi-party data sharing	54
4.2. Public investment and institutional design	56
4.3. Standards, and cross-border interoperability and convergence	60
4.4. Certification and public trust frameworks	62
4.5. Inter-governmental (G2G) data spaces and government as operator and anchor participant	62
References	66
Endnotes	69

5 Conclusions and policy considerations 70

5.1. Targeting investment where the structural condition holds	71
5.2. Distinguishing between funding for the technical foundation and for ecosystems	71
5.3. Addressing the demand-side funding gap	72
5.4. Supporting business models around infrastructural value	72
5.5. Supporting adoption of trust-enhancing technologies and their integration by design	73
5.6. Adopting certification and government anchoring to build institutional trust	73
5.7. Recognising AI-driven data demand as a pathway to viability	74
5.8. Enhancing measurement	74
References	75

Annex A. OECD Workshop on “Assessing the Sustainability of Business Models in Data Spaces and the Role of Public Policies” 76

FIGURES

Figure 1.1. Development stage distribution of 235 data space initiatives	13
Figure 2.1. Stylised architecture of a data space	23
Figure 3.1. Building block adoption rates: All entries vs. operational data spaces	32
Figure 3.2. Operational share by sector	34
Figure 4.1. EU public funding rate by development stage	58

TABLES

Table 3.1. Comparison of data-sharing mechanisms and enablers	33
Table 3.2. Governance configurations of data spaces based on type of actors	44

BOXES

Box 1.1. The evidence base: Data source limitations and the measurement gap	14
Box 2.1. The landscape of data spaces standardisation	21
Box 2.2. SDMX: A mature data-sharing ecosystem for official statistics	23
Box 2.3. The “data fabric”: Data mesh extended to inter-organisational settings	25
Box 3.1. The case of the Catena-X Automotive Network	36
Box 3.2. Privacy-enhancing technologies as complement to contract-based usage control	40
Box 3.3. The case of the Maritime Data Space	42
Box 3.4. The case of Agdatahub	43
Box 3.5. Telecommunications operators as data space service providers: The case of T-Systems	45
Box 4.1. The case of the Automotive and Battery Traceability Centre (ABtC)	56
Box 4.2. Inter-governmental (G2G) data-sharing frameworks in the public sector	64

Abbreviations and acronyms

AA	Account Aggregator
ABtC	Automotive and Battery Traceability Centre
AI	Artificial intelligence
CESMII	Clean Energy Smart Manufacturing Innovation Institute
DEPA	Data Empowerment and Protection Architecture
DFFT	Data free flow with trust
DPPS	Data Policy, Privacy and Online Safety
DGA	Data Governance Act
DMREF	Designing Materials to Revolutionize and Engineer Our Future
DPC	Digital Policy Committee
DSSC	Data Spaces Support Centre
EHDS	European Health Data Space
EU	European Union
ICIS	International Conference on Information Systems
IDS	Industrial Data Space
IDSA	International Data Spaces Association
IEEE	Institute of Electrical and Electronics Engineers
IOFDS	International Open Forum on Data Society
IPA	Information-technology Promotion Agency
MDS	Maritime Data Space
METI	Ministry of Economy, Trade and Industry
ODS	Open Data Spaces
OEM	Original equipment manufacturer
PET	Privacy-enhancing technologies
PPDS	Public Procurement Data Space
R&D	Research and development
RAM	Reference Architecture Model
SME	Small and medium-sized enterprise
TEFCA	Trusted Exchange Framework and Common Agreement

Executive summary

Findings

Growing demand for high-quality data is raising the cost of barriers to data sharing across borders.

Sharing data across organisations and jurisdictions generates economic and social value – from improving supply chain resilience to accelerating healthcare research. Yet many organisations that hold valuable data often lack the incentives, governance frameworks or legal certainty to share them. Sharing data has long been an economic and social necessity, but it has intensified as progress in artificial intelligence (AI) from scaling computation alone appears to slow, and increasingly depends on access to high-quality, domain-specific data. Centralised commercial platforms have partially closed this gap, but they require participants to trust a single data intermediary with their data, a condition that is not feasible in all settings.

Data spaces have emerged as a response to this type of challenge, and functionally equivalent approaches exist globally.

A data space is a data ecosystem characterised by (i) federated or decentralised governance; (ii) open, standards-based interoperability; and (iii) mutual trust and effective control between data providers and data users over how data are used. Policy efforts on data spaces are often associated with European initiatives, where public investment in data spaces exceeded EUR 1.7 billion as of 2025. However, functionally equivalent approaches, sometimes developed independently and under different terminologies, have emerged. Key examples can be found in Japan (Open Data Spaces, ODS), the United States (Trusted Exchange Framework and Common Agreement, TEFA) and India (Data Empowerment and Protection Architecture, DEPA). This suggests that multi-party data sharing under federated governance addresses a structural need that is not specific to any single jurisdiction or technical architecture.

Data spaces are best suited for a specific class of use cases involving multi-party data sharing.

The analysis identifies four conditions that, taken together, distinguish settings where data spaces offer advantages over common alternatives, such as bilateral contracts, centralised platforms, privacy-enhancing technologies (PETs): (i) data sharing among three or more parties; (ii) competitive or jurisdictional relationships that preclude trusting a single data intermediary; (iii) non-excludable collective value that bilateral contracts alone cannot capture; and (iv) ongoing data flows requiring continuous governance. These features position data spaces as key infrastructures for advancing data free flow with trust (DFFT), as demonstrated by operational use cases in mobility, maritime logistics and supply chain management, where the cumulative conditions of competition, regulatory pressure and cross-border collaboration apply.

Despite significant investment, few data spaces have reached financial and operational viability.

Case study evidence from industry-led Catena-X, the Finnish Maritime Data Space (MDS), and Japan's Automotive and Battery Traceability Centre (ABtC) show that data spaces work in practice. The International Data Spaces Association (IDSA) Radar, the largest publicly available registry of data space initiatives, lists 235 entries as of February 2026, nearly triple the 87 recorded in 2022. European initiatives dominate the listings, reflecting the historical concentration of data space activity around EU-funded programmes. Only 25 of these initiatives (one in nine) have reached operational status, and none have achieved the scaling stage associated with financial viability. Most depend on public funding, and few have business models that could sustain operations independently. Once operating at scale, data spaces can more systematically support trusted data sharing across jurisdictions. This has already been demonstrated as a proof of concept by the exchange of supply chain data between Catena-X and the Ouranos Ecosystem across the European Union and Japan.

Establishing a data space faces a bootstrapping problem. The shared infrastructure must precede the network effects that justify it. Moreover, each participant's onboarding investment must be justified by returns that may not materialise until the ecosystem reaches sufficient scale. This chicken-and-egg dynamic, referred to as the “bootstrapping problem”, is a central barrier to data space viability. In cases where it has been overcome, as in the Finnish MDS, quantifiable use-case value was present from the outset. ABtC in Japan followed the same pattern: original equipment manufacturers and suppliers were anchored by a concrete initial use case (carbon footprint data) before the scope expanded across the battery lifecycle. Where the bootstrapping problem has not been overcome, as with Agdatahub (which declared bankruptcy in 2024), infrastructure was built without demonstrated use-case value sufficient to sustain participation once public grant funding expired.

The most limiting factor may be institutional design rather than funding. Continued funding may itself reduce the pressure on data space ecosystems to develop business models that can sustain them after the subsidy ends. Where funding rewards grant compliance over outcomes, rational actors optimise for the former, a misaligned incentive documented in public venture capital. Overlapping initiatives developing redundant building blocks compound this risk in two ways: duplicative research and development (R&D); and delays driven by possible vested interests, another pattern well documented in technology standardisation more broadly.

Conclusions

Based on the analysis above, eight policy considerations emerge for supporting the adoption of viable data spaces:

- **Target investment where the structural condition holds.** A two-stage assessment could help direct public resources more effectively towards use cases where data spaces are both structurally necessary and practically viable. The first stage needs to test whether the co-ordination problem meets the four conditions highlighted above. The second needs to assess whether economic and sectoral preconditions are in place: pre-existing digital infrastructure among prospective participants, identifiable anchor firms and clear value propositions.
- **Distinguish between funding the technical foundation and funding ecosystems.** Investment in information models, protocols and interoperability frameworks constitutes pre-competitive R&D with broad spillover benefits. Investment in ecosystem creation depends on whether a specific collective-action problem has been identified that no market actor has the incentive or legitimacy to solve. Separating these purposes in policy design would allow better calibrated interventions.
- **Address the demand-side adoption gap.** The dominant policy response has been supply-side funding directed at shared infrastructure. Far less common are measures that reduce the onboarding costs individual participants face. Governments may wish to complement infrastructure funding with demand-side measures. This could be direct subsidy for participant integration (as in Spain's Data Spaces Kit), capacity building (as in the educational programme of Japan's Information-technology Promotion Agency) or regulatory design that makes the intermediary role commercially viable (as in India's DEPA).
- **Support business models around infrastructural value.** Viable models in operational data spaces resemble cost-sharing arrangements and membership-based infrastructure funding, combined with value-added services rather than data monetisation. Prioritising use-case value over infrastructure development is consistently associated with stronger prospects for financial viability. Where the value generated is collective and infrastructural in character, public support may be justified.

- **Support adoption of trust-enhancing technologies and their integration by design.** Data usage control (built on dataspace protocols) and PETs (e.g. federated learning, secure multi-party computation) offer complementary guarantees. Sustained support for both, and for their interoperability, would promote their respective adoption and strengthen trust among participants.
- **Adopt certification and government anchoring to build institutional trust.** In sectors where no natural anchor exists, governments may wish to consider designating institutional anchors, including public bodies or state-owned enterprises. Governments can also lead by example by deploying data spaces for their own operations, including inter-governmental (G2G) configurations. In this way, it would demonstrate that the architecture works at scale, lowering the perceived adoption risk for private participants. Certification mechanisms can complement this by generating trust signals to prospective participants. Estonia's X-Road and Japan's "public interest digital platform operator" certification offer promising models.
- **Recognise AI-driven data demand as a pathway to viability.** The demand for high-quality, domain-specific training data creates business incentives for governed data sharing that previous waves of data space formation may have lacked. Whether AI-driven demand proves transformative depends on whether data spaces can enable data sharing at the speed and scale that AI system development and use require.
- **Enhance measurement.** Few data space initiatives collect the data needed to evaluate whether they are delivering value that alternative mechanisms could not. The measurement problem has structural roots: the federated nature of data spaces makes performance tracking difficult as typically no single entity observes all transactions. Extending systematic data collection across OECD Members and partners, including to non-European federated data-sharing initiatives, is a precondition for evidence-based policy guidance that is genuinely international in scope.

1 Overview

Data sharing across organisations and jurisdictions generates economic and social value, increasingly so as progress in artificial intelligence depends on high-quality, domain-specific data. This chapter sets out why data spaces have emerged as a response to the trust and collective-action problems that centralised platforms leave unresolved, and why few of the more than 200 registered initiatives have reached operational viability despite significant public investment. It frames the question the report addresses, namely under what conditions data spaces provide value that alternative data-sharing mechanisms cannot, and describes the evidence base and its limitations.

Data sharing generates economic value through multiple channels: it strengthens supply chain resilience by enabling firms to anticipate disruptions; raises productivity by reducing duplicative data collection; accelerates innovation by allowing novel combinations of datasets that no single organisation could assemble alone; and more generally provides the diverse, domain-specific training data on which artificial intelligence (AI) systems increasingly depend (OECD, 2019^[1]; Jones and Tonetti, 2020^[2]; OECD, 2025^[3]). These benefits are well established. Indeed, the policy case for enhancing data access and sharing has been articulated across successive OECD legal instruments, most recently the OECD Recommendation on Enhancing Access to and Sharing of Data [\[OECD/LEGAL/0463\]](#) (OECD, n.d.^[4]).

The urgency of this agenda is sharpening. As the pace of AI capability gains from scaling training data and compute alone appears to be slowing down, attention is shifting towards data quality, curation, and domain specificity as the binding constraints on further progress (Hoffmann et al., 2022^[5]; Zeff, 2024^[6]; Lu, 2025^[7]). Publicly available data from the open web may be abundant, but they are often noisy, biased or legally contested. Organisations already exploit their own operational records, sensor streams, clinical observations and supply chain metrics for internal AI applications, including through retrieval-augmented generation and domain-specific model fine-tuning. What they lack are the governance frameworks, legal certainty and control mechanisms to reduce the risks of data sharing, and often a clear view of the economic returns. This is particularly true where combining complementary datasets from multiple holders could unlock substantially greater value, but that value accrues collectively rather than to any single participant.

The commercial data platform market provides supportive evidence of this economic value. Combined revenues of two of the leading data sharing platforms exceeded USD 10 billion in 2025, serving over 30 000 customers including most of the Fortune 500 firms.¹ These platforms operate data marketplaces and cross-organisation exchange protocols that reach thousands of data recipients across competing ecosystems. The commercial data platform market has demonstrated that organisations will share data when the infrastructure is convenient, the network effects are strong and the value proposition is immediate.

Yet certain use cases impose conditions that centralised platforms may be unable to meet. For instance, firms will not share competitively sensitive data (e.g. supply chain positions, production quality metrics, emissions profiles) unless they trust that usage conditions will be enforced and that governance arrangements do not unfairly advantage one participant over others. For many use cases, these limitations may not be significant: the platform's functionality, speed and network effects might outweigh the governance constraints. For others, however, the limitations may be binding constraints. This may be especially true where the value of shared data is collective and infrastructural in character, generating benefits that extend beyond any single transaction and organisation (Frischmann, 2012^[8]). The questions become: which problems fall into each category, and is the distinction sufficiently sharp to guide investment and public policies?

1.1. Objective: Assessing the distinctive value of data spaces

Data spaces have emerged as one of the principal data governance responses to the trust and collective-action problems that centralised platforms leave unresolved. These collective-action problems, which involve multi-party data sharing among possible competitors, across jurisdictions and under data control requirements, are not specific to any one region. They arise wherever digitalisation, supply chain integration or AI development require controlled access to distributed data. The largest concentration of data space initiatives has emerged within the European policy context. This has been driven by the European Strategy for Data (European Commission, 2020^[9]), the EU Data Governance Act (European Union, 2022^[10]), the EU Data Act (European Union, 2023^[11]) and associated public investment

programmes (see Chapter 4). However, data spaces are gaining policy recognition and adoption around the world, sometimes via functionally equivalent arrangements under different names (see Chapter 2).

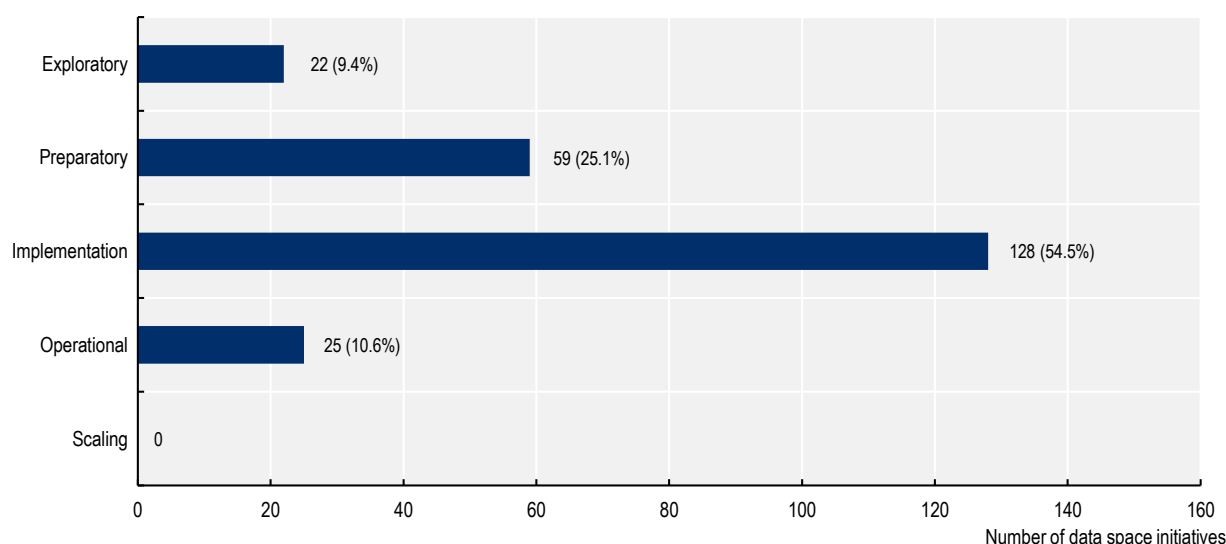
Outside Europe, Japan has advanced the furthest. The Ouranos Ecosystem, an industrial initiative initiated and supported by Japan's Ministry of Economy, Trade and Industry (METI), draws on the Open Data Spaces (ODS) technical architecture developed by Japan's Information-technology Promotion Agency (IPA). It has produced operational deployments such as the Automotive and Battery Traceability Centre (ABtC) (see Chapters 2 and 3). Other countries such as Canada and Korea, as well as Brazil and the People's Republic of China (hereafter "China"), are at earlier or accelerating stages. Outside the data spaces label, India's Data Empowerment and Protection Architecture (DEPA) and the Trusted Exchange Framework and Common Agreement (TEFCA) in the United States for health data are examples of functionally equivalent arrangements.

More recently, the United States launched the Genesis Mission in November 2025 (The White House, 2025^[12]). This directs the creation of a federated data-sharing platform across the 17 US national laboratories to accelerate AI-enabled research in priority sectors, including nuclear fusion, quantum computing, advanced manufacturing, energy infrastructure and critical materials. Initial investments by the US Department of Energy include approximately USD 320 million announced in December 2025 to build AI infrastructure (the American Science Cloud and the Transformational AI Models Consortium). This was followed by a USD 293 million Request for Applications in March 2026 inviting interdisciplinary teams to address the programme's identified national science and technology challenges (US Department of Energy, 2025^[13]; US Department of Energy, 2026^[14]). While the programme does not adopt the data spaces label, it functionally creates a federated data-sharing model. This, in turn, highlights the importance of intellectual property frameworks for contributed proprietary training data, an issue addressed in Section 3.3 of Chapter 3.

The International Data Spaces Association (IDSA) Radar,² the largest publicly available registry of data space initiatives, lists 235 entries as of February 2026 across sectors ranging from mobility and health to agriculture and energy. The IDSA Radar nearly tripled in size from 87 entries in 2022 to 235 in early 2026, confirming wide-scale experimentation and growing maturity of data spaces. Yet the ecosystem remains overwhelmingly pre-operational. Of the 235 initiatives, only 25 have reached operational status. None have achieved what the IDSA Radar classifies as "scaling", the only stage that IDSA's definitions explicitly associate with financial and operational viability (IDSA, 2024^[15]) (Figure 1.1).³ Only one in nine entries has reached operational status, a "valley of death"⁴ that constitutes the defining challenge facing the ecosystem (Box 1.1). This warrants particular attention given that public investment in data space initiatives in the European Union alone is estimated at over EUR 1.7 billion as of early 2025 (Otto and Tardieu, 2025^[16]).⁵

This gap between investment and operational maturity has raised a pointed question among policymakers around the world: are data spaces worth the investment? This paper reframes that question in more precise terms. The issue is not whether data spaces create value in the generic sense that any data-sharing infrastructure creates value; commercial platforms have already demonstrated that proposition. Rather, the question is whether certain data-sharing problems require the specific properties that data spaces provide, in particular federated architecture, provider-defined usage control, distributed governance and open interoperability. If such a class of problems exists, then the gap between investment and operational scaling identified above is worth closing, and policy attention may need to focus on how to close it most effectively. If it does not, then other approaches including centralised platforms are the rational default, and the policies directed towards data spaces would deserve re-evaluation.

Figure 1.1. Development stage distribution of 235 data space initiatives



Note: Development stages are self-reported by data space operators to the IDSA Radar and may not be updated as initiatives progress or fail. Entries are predominantly European in origin, reflecting the historical concentration of data space activity around EU-funded programmes (see Box 1.1). “Implementation” denotes a detailed project plan and resources for a pilot but no operational use case. “Operational” denotes at least one use case with data flowing between providers and recipients. One entry classified as “Scaling” on the IDSA Radar dashboard is recorded without a development stage in the underlying dataset export; the figure therefore shows zero scaling entries.

Source: Based on IDSA (n.d.), *IDSA Data Spaces Radar*, www.dataspaces-radar.org.

1.2. Scope and approach

This paper provides an evidence-based assessment of the distinctive value proposition of data spaces. In particular, it asks under what conditions, and through what mechanisms, do data spaces solve problems that alternative data-sharing approaches cannot.

This focus on value precedes, and is a necessary condition for, the question of business models. Business model analysis asks how an organisation creates, delivers and captures value (Osterwalder and Pigneur, 2010^[17]). As applied to data spaces, authors such as Mertens et al. (2024^[18]) decompose this into the use cases that generate value; the value proposition offered to participants; the implementation model; and the balance between cost structure and revenue. However, this decomposition presupposes that the value in question has been identified. If it is misidentified, business models built on that assumption risk being structurally misaligned with the value the data space may produce. A data space whose value is collective and infrastructural, for instance, may not be able to sustain business models designed around transactional data monetisation.

IDSA Radar data are consistent with this pattern: data spaces that adopted marketplace-based revenue models tend to struggle to reach higher degrees of maturity. That said, this should be read as indicative rather than conclusive given the relatively small sample of operational data spaces (25 entries) and their self-reporting nature (Box 1.1). The present paper therefore addresses two questions: what distinctive value do data spaces create, and through which mechanisms? Only once this foundation is established can the question of how to capture that value through viable business models be answered soundly.

To mitigate the data source limitations highlighted in Box 1.1, the empirical base of this draft report combines the quantitative data from the IDSA Radar with case study evidence from, and international comparisons with, parallel developments in Japan (ABtC, the ODS ecosystem), India (DEPA), and the United States (TEFCA for health data). The analysis also draws on findings from the OECD Workshop on Assessing the Sustainability of Business Models in Data Spaces and the Role of Public Policies (hereafter

“the Expert Workshop”). Held on 11 March 2025 at the Data Spaces Symposium in Warsaw, the workshop brought together practitioners, policymakers and industry experts from Europe, Japan and North America. Together, they critically examined the financial sustainability of data space business models on the basis of concrete use cases, both successful and failed (Annex A).⁶

Box 1.1. The evidence base: Data source limitations and the measurement gap

The IDSA Radar is a voluntary, self-reported registry whose registered initiatives are predominantly European, reflecting both a registration bias and the European concentration of data space activity noted in Section 1.1. Where this paper draws on IDSA Radar statistics, the quantitative patterns describe the European data spaces ecosystem as registered in the IDSA Radar, not necessarily the global landscape of federated data-sharing arrangements.

Beyond this coverage bias, the Radar presents other specific data quality constraints. Missing data rates are high: approximately 50% of entries lack geographical focus, 58% lack funding source information and 78% lack start dates. Development stages are recorded at a point in time and may not reflect current status; initiatives that have failed may still appear as operational. Analyses of subsets (funding, governance, access models) are based on the reporting population only.

This measurement gap is itself a key finding. Most operational data spaces do not track the metrics needed to assess their own value. The Radar captures no outcome data: no transaction volumes, no participant growth trajectories, no cost comparisons, no productivity indicators, and no measures of data flows enabled or withheld. Case studies provide qualitative evidence of mechanism activation but were not designed as controlled evaluations. Consequently, they cannot isolate the causal contribution of data space architecture from confounding factors such as regulatory incentives, pre-existing industry relationships or the maturity of participating firms’ digital infrastructure.

The measurement problem has structural roots. Data spaces are, by design, distributed systems in which no single entity observes all transactions. The very data provider control principles that define data spaces make centralised data collection difficult. Participants may be reluctant to share performance data with a registry for the same reasons they are reluctant to share operational data with centralised platforms. The result is a paradox: the data-sharing infrastructure that data spaces provide is itself under-measured because the principles that make it distinctive also make it hard to observe.

A deeper issue compounds this difficulty: Measuring the value of data. The value of data is not intrinsic: it depends on context, governance, and the extent to which it is shared and linked with complementary datasets. No single valuation method captures the full picture (OECD, 2022^[19]; Bisière et al., 2025^[20]). In particular, data use and sharing generate spillovers, both positive and negative, that are not reflected in market prices. These challenges apply even more severely to data spaces, where value is predominantly collective and infrastructural (Chapter 3): avoided costs, system-level resilience, ecosystem-wide regulatory compliance and innovation spillovers from unanticipated data combinations all resist conventional quantification. The proposed DGP Measurement Agenda [DSTI/DPC/DGP(2026)4] aims to address this broader gap.

The paper is structured as follows. Chapter 2 establishes a two-tier working definition for data spaces that distinguishes broad (architecture-agnostic) from narrow (federated, with provider control) formulations. Chapter 3 identifies the distinctive value proposition of data spaces by analysing the co-ordination problem they address; the boundary conditions under which, and operational mechanisms through which, they hold a comparative advantage over alternative data-sharing mechanisms; and the bootstrapping challenge that explains the maturity and adoption gap to a significant extent. Chapter 4 provides a brief overview of the emerging policy landscape, examining regulatory drivers, public investment design and cross-border interoperability efforts across jurisdictions. Chapter 5 draws conclusions and sets out policy considerations related to data space initiatives.

References

- Bisière, C. et al. (2025), “The economics of data spaces”, *Policy Paper*, SE Digital Center and Toulouse School of Economics, France, https://www.tse-fr.eu/sites/default/files/TSE/documents/DigitalCenter/policy_paper/the_economics_of_data_spaces_july_2025_policy-paper.pdf. [20]
- Data Spaces Symposium (2025), “Memories of Data Spaces Symposium 2025”, <https://www.data-spaces-symposium.eu/program-2025> (accessed on 1 March 2026). [26]
- Databricks (2026), “Databricks grows >65% YoY, surpasses \$5.4 billion revenue run-rate”, <https://www.databricks.com/company/newsroom/press-releases/databricks-grows-65-yoy-surpasses-5-4-billion-revenue-run-rate> (accessed on 9 February 2026). [25]
- European Commission (2020), *Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the regions: A European strategy for data*, European Commission, Brussels, <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52020DC0066>. [9]
- European Union (2023), *Regulation (EU) 2023/2854 on harmonised rules on fair access to and use of data (Data Act)*, European Union, Brussels. [11]
- European Union (2022), *Regulation (EU) 2022/868 of the European Parliament and of the Council of 30 May 2022 on European data governance and amending Regulation (EU) 2018/1724 (Data Governance Act)*, European Union, Brussels, <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32022R0868>. [10]
- Frank, C. et al. (1996), “Surviving the “valley of death”: A comparative analysis”, *The Journal of Technology Transfer*, Vol. 21, pp. 61-69, <https://doi.org/10.1007/bf02220308>. [24]
- Frischmann, B. (2012), *Infrastructure: The Social Value of Shared Resources*, Oxford University Press. [8]
- Hoffmann, J. et al. (2022), “Training compute-optimal large language models”, *arXiv*, Vol. 35/2203.15556v1, <https://arxiv.org/abs/2203.15556v1>. [5]
- IDSA (2024), *IDS Reference Architecture Model, Version 4.0*, https://github.com/International-Data-Spaces-Association/IDS-RAM_4_0. [15]
- IDSA (n.d.), “International Data Spaces”, <https://internationaldataspaces.org/> (accessed on 31 March 2026). [23]
- Jones, C. and C. Tonetti (2020), “Nonrivalry and the economics of data”, *American Economic Review*, Vol. 110/9, pp. 2819-2858, <https://doi.org/10.1257/aer.20191330>. [2]
- Lu, C. (2025), “The race to efficiency: A new perspective on AI scaling laws”, *arXiv* 2501.02156v3, <https://arxiv.org/abs/2501.02156v3>. [7]
- Markham, S. et al. (2010), “The valley of death as context for role theory in product innovation”, *Journal of Product Innovation Management*, Vol. 27/3, pp. 402-417, <https://doi.org/10.1111/j.1540-5885.2010.00724.x>. [22]
- Mertens, C. et al. (2024), “Data spaces business models”, *Position Paper*, Zenodo (CERN European Organization for Nuclear Research), <https://doi.org/10.5281/zenodo.14101303>. [18]

- OECD (2025), *Enhancing Access to and Sharing of Data in the Age of Artificial Intelligence*, OECD Publishing, Paris, <https://doi.org/10.1787/23a70dca-en>. [3]
- OECD (2022), “Measuring the value of data and data flows”, *OECD Digital Economy Papers*, No. 345, OECD Publishing, Paris, <https://doi.org/10.1787/923230a6-en>. [19]
- OECD (2019), *Enhancing Access to and Sharing of Data: Reconciling Risks and Benefits for Data Re-use across Societies*, OECD Publishing, Paris, <https://doi.org/10.1787/276aaca8-en>. [1]
- OECD (2013), *Commercialising Public Research: New Trends and Strategies*, OECD Publishing, Paris, <https://doi.org/10.1787/9789264193321-en>. [27]
- OECD (n.d.), *Recommendation of the Council on Enhancing Access to and Sharing of Data*, OECD/LEGAL/0463, <https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0463> (accessed on 6 March 2023). [4]
- Osterwalder, A. and Y. Pigneur (2010), *Business Model Generation: A Handbook for Visionaries, Game Changers, and Challengers*, Wiley, New York. [17]
- Otto, B. and H. Tardieu (2025), *Data Sharing: Closing the Gap in the Draghi Report’s Call for AI Competitiveness in Europe*, Le Grand Continent. [16]
- Snowflake (2026), “Financial results for the fourth quarter and full-year of fiscal 2026”, <https://www.snowflake.com/en/news/press-releases/snowflake-reports-financial-results-for-the-fourth-quarter-and-full-year-of-fiscal-2026/> (accessed on 28 April 2026). [21]
- The White House (2025), “Launching the Genesis mission”, <https://www.whitehouse.gov/presidential-actions/2025/11/launching-the-genesis-mission/> (accessed on 28 April 2026). [12]
- US Department of Energy (2026), “Energy department announces \$293 million in funding to support genesis mission national science and technology challenges”, <https://www.energy.gov/articles/energy-department-announces-293-million-funding-support-genesis-mission-national-science> (accessed on 28 April 2026). [14]
- US Department of Energy (2025), “Energy department advances investments in AI for science”, <https://www.energy.gov/articles/energy-department-advances-investments-ai-science> (accessed on 28 April 2026). [13]
- Zeff, M. (2024), “Current AI scaling laws are showing diminishing returns, forcing AI labs to change course”, *TechCrunch*, <https://techcrunch.com/2024/11/20/ai-scaling-laws-are-showing-diminishing-returns-forcing-ai-labs-to-change-course/> (accessed on 16 March 2026). [6]

Endnotes

¹ Snowflake reported USD 4.7 billion in revenue for fiscal year 2026, ending 31 January 2026 (Snowflake, 2026^[21]), while Databricks reached an annualised revenue of USD 5.4 billion as of January 2026 (Databricks, 2026^[25]).

² The IDSA Radar is a self-reported, voluntary registry maintained by the International Data Spaces Association (IDSA), a not-for-profit association based in Germany with more than 140 member organisations, which “brings together global members from both industry and research” to “develop and promote the concept of data spaces” (IDSA, n.d.^[23]).

³ Over half of all entries (128, or 54%) are at the implementation stage (project plan and resources but no operational use case). Only 25 (11%) have reached operational status. Public funding dominates the early stages (46.1% of implementation-stage data spaces received EU funding). However, it recedes sharply at maturity (20% of operational ones), and no market-based revenue model has taken its place: only 2 of 96 initiatives reporting an access model charge a fee. The ecosystem is not merely young; it faces a structural viability gap.

⁴ The “valley of death” refers to the critical gap between early-stage development and self-sustaining operation, during which initiatives are most vulnerable to failure. The concept originates in technology commercialisation literature, where it describes the funding and capability gap between research output and market viability (Frank et al., 1996^[24]; Markham et al., 2010^[22]). The OECD has applied the concept to science-based innovation policy, identifying the valley of death as the phase where public research funding has ended but private investment has not yet materialised (OECD, 2013^[27]).

⁵ The EUR 1.7 billion figure (EUR 1 735.4 million) originates from a funding portfolio compiled by the Gaia-X European Association for Data and Cloud AISBL and first published in Otto and Tardieu (2025), a policy commentary in *Le Grand Continent*. Both authors hold or held board positions at Gaia-X AISBL and IDSA. The compilation aggregates EU-level programmes (Digital Europe Programme work programmes 2021-2022 and 2023-2024, EU4Health, Horizon Europe and Destination Earth) with national funding from six member states (Germany, Spain, France, Luxembourg, Denmark and Finland). It includes programmes whose scope extends beyond data spaces as narrowly defined, notably Destination Earth (EUR 240 million) and a broad EU4Health allocation (EUR 280 million). It omits known national investments from other member states. No published methodology accompanies the compilation. The figure should therefore be read as an indicative order of magnitude rather than a precise audited total; the actual aggregate, once additional member state investments are included, is likely higher.

⁶ The workshop was governed by the Chatham House Rule. A summary of workshop discussions is provided in Annex A. For the agenda, see Data Spaces Symposium (2025^[26]).

2 Understanding data spaces

This chapter traces how the data space concept evolved from a computer science approach to data integration into a governance framework for trusted multi-party data sharing. It reviews emerging definitions and reference architectures and their convergence on data provider control. It then proposes a two-tier working definition, distinguishing a broad, architecture-agnostic formulation from a narrow one centred on federated governance, open interoperability and provider control, and uses this definition to situate data spaces relative to adjacent concepts and to functionally equivalent initiatives outside Europe.

“Data space” has meant different things at different times. What began as a computer science abstraction now denotes a governance framework for multi-party data sharing backed by an estimated EUR 1.7 billion in public investment in the European Union (EU) alone. This chapter traces how the concept shifted; what the current formulation adds to earlier data-sharing approaches; and how data related to other comparable concepts.

2.1. From data integration to data governance

The term “dataspace” originated in computer science as a response to a practical problem: traditional data integration requires complete schema alignment before any useful service can be provided. This imposes costs that scale with every new connection between semantically heterogeneous and administratively autonomous sources. Franklin, Halevy and Maier (2005^[1]) proposed data spaces as an alternative paradigm that inverts this sequence. Rather than requiring upfront integration, a data space provides basic services immediately (cataloguing, search, browsing) across all participating sources, adds approximate query answering with minimal mappings, and invests in full semantic alignment only where the return justifies the cost. They called this “pay-as-you-go” integration: each level of integration delivers value, and organisations need not achieve full interoperability before extracting benefit.¹

The Industrial Data Space (IDS), proposed by the Fraunhofer Institutes and Research Units and formalised in a Reference Architecture Model (RAM) (Otto et al., 2017^[2]), was a significant shift from the previous paradigm. IDS transformed the dataspace concept by embedding governance and trust directly into the architecture. ETSI’s recent European standard for a data catalogue framework (EN 304 199) captures this shift in definitional form. It characterises a data space as an “environment enabling trusted data sharing between participating parties, based on an agreed governance framework, along with an agreed set of policies, semantic models, standardised protocols, processes and facilitating services” (ETSI, 2025^[3]).

Two features mark the transition:

1. *IDS adopts a federated approach:*² data remain with the provider, and access and use are mediated by certified connectors (i.e. software components deployed by each participant that govern what data are shared, with whom and under what conditions). Meanwhile, shared services (identity management, metadata brokering, certification) operate at the ecosystem level.
2. *IDS introduces data usage control:* the conditions under which connectors grant access are not merely contractual but machine-readable. That means data providers can attach usage rules specifying how shared data may be processed, aggregated and forwarded, and connectors can enforce these rules at the point of access. This combination of federated custody and enforceable usage control reflects a fundamentally different concern from the original dataspace concept: not how to query heterogeneous data, but how to enable its flow across organisational boundaries while still preserving data control among autonomous parties (Otto et al., 2017^[2]).

More than one architectural approach now uses the term “data space”. IDS, and the ecosystem that developed around it (Eclipse Dataspace Protocol, European Data Spaces Support Centre Blueprint), prioritises the governance problem. As such, it focuses on embedding trust, identity and usage control via machine-readable contract negotiation into a connector architecture that presupposes co-ordinated adoption of a shared protocol stack. ODS, a separate architecture developed by the Information-technology Promotion Agency (IPA) (METI/IPA, 2025^[4]; IPA, 2026^[5]), prioritises architectural decoupling. IPA describes ODS as “an open and scalable foundation for distributed data management, built on organisational and national diversity by design” (IPA, 2026^[5]). ODS draws its data management architecture from two sources: the original dataspace concept of Franklin, Halevy and Maier (2005^[1]); and Data Mesh principles (domain ownership, data-as-a-product, federated governance) that had been developed by Dehghani (2022^[6]) for intra-organisational settings.

IPA organises its reference architecture around three loosely coupled pillars that can be independently replaced or upgraded (ontology and semantic interoperability; data addressability and discoverability; and identity and usage control; see Section 3.3.2 of Chapter 3). For its part, ODS engages with governance concepts that IDS had formalised, including usage control (Steinbuss et al., 2021^[7]) and trust, and maintains a degree of logical compatibility with the IDS-RAM (METI/IPA, 2025^[4]). However, it makes different architectural choices about how to implement them. Both approaches target inter-organisational and cross-border scope and diagnose the same governance challenges, but they differ in architectural foundations and in what they assume can be co-ordinated *ex ante* (Section 3.3 of Chapter 3).

Usage control, however technically enforced, addresses how shared data may be processed but does not resolve who owns the value derived from them. Where data spaces are used to pool proprietary data for AI model training, this surfaces a distinct governance question: how to attribute ownership of resulting models, derived insights or downstream patent claims among data contributors, model developers and platform operators. Public-sector experience with multi-party data sharing for research and development includes the Materials Genome Initiative in the United States. The initiative has committed more than USD 1 billion in cumulative federal investment (NSF, 2023^[8]). However, it has encountered persistent industry reluctance to contribute proprietary performance and testing data, which illustrates the practical weight of this question. Approaches such as privacy-enhancing technologies (PETs) for anonymised contribution and federated learning, as well as revenue-sharing coalitions, have addressed parts of the problem. Still, none has yet produced a widely adopted commercial model. As data spaces increasingly position themselves as substrates for AI training (Section 4.2.2 of Chapter 4), the absence of a clear framework for attributing intellectual property rights risks becoming a key constraint on participation, particularly for holders of high-value proprietary data.

2.2. Emerging definitions and converging architectures

No single universally agreed definition of “data space” exists. Instead, multiple institutional actors are arriving at overlapping definitions through different institutional pathways. ISO/IEC DIS 20151, the most broadly scoped definition under development, defines a data space as a “governance framework and supporting services to build trustworthiness and enable data sharing through an agreed set of policies, semantic models, protocols and processes” (ISO, 2025^[9]). (See Box 2.1) This formulation is deliberately architecture-agnostic: it specifies neither centralisation nor decentralisation, neither a particular topology nor a distribution model. The European Data Spaces Support Centre (DSSC) adopts a similar architecturally neutral definition, converging with the ISO standard.

Definitions that embed specific structural requirements occupy the other end of this spectrum. For instance, IDSA defines a data space as “the total set of interoperable data-sharing applications by actors in a specific sector or domain [...] based on a distributed software infrastructure which provides the required software functionality for data sovereignty and trust” (IDSA, 2024^[10]). This definition embeds two structural requirements absent from the ISO formulation: (i) a distributed infrastructure; and (ii) technically enforced “data sovereignty”.

The International Open Forum on Data Society (IOFDS) recommends a further definition. A standing forum convened by Japan’s Data Society Alliance, with participation from IDSA, Gaia-X, FIWARE and the Institute of Electrical and Electronics Engineers (IEEE), IOFDS advances interoperability between data space initiatives across regions. IOFDS defines a data space as “a decentralized ecosystem with common policy and rules defined by a governance framework that enables secure and trustworthy data transactions between participants while supporting trust and data sovereignty” (IOFDS, n.d.^[11]).³

Despite their differences, these definitions converge on what the data spaces community calls “data sovereignty”. In this context, data sovereignty is understood as the ability of data providers to retain effective control over how their data are accessed and used after sharing.⁴ This shared commitment is

operationalised through three major reference architectures (the IDSA-RAM, the ODS-RAM and the DSSC Blueprint), each reflecting a different model of public-private collaboration (Section 4.3 of Chapter 4).

Convergence between the three reference architectures is progressing. In March 2025, for example, IPA and Catena-X completed a cross-border interoperability proof of concept. This demonstrated product carbon footprint data exchange between the Ouranos Ecosystem and Catena-X for EU Battery Regulation compliance (IPA and Catena-X, 2025^[12]) (Section 4.3 of Chapter 4). The trend towards convergence is being formalised through an active, multi-layered standardisation process (Box 2.1).

Box 2.1. The landscape of data spaces standardisation

Four standardisation bodies are developing complementary standards for data spaces. At the international level, ISO JTC1 SC38 (Cloud Computing and Distributed Platforms) has three draft international standards: ISO/IEC DIS 20151-1 (dataspace concepts and characteristics); ISO/IEC DIS 26450 (Eclipse Dataspace Protocol, DSP); and ISO/IEC DIS 26451 (Eclipse Decentralised Claims Protocol, DCP). Upcoming work items include dataspace trust frameworks (ISO/IEC 20151-2), use cases (ISO/IEC 25850), switching by design and observability. The Eclipse Dataspace Protocol and Decentralised Claims Protocol appear in both the ISO and CEN/CENELEC pipelines. This positions them as the de facto technical baseline, with formal de jure status pending ISO/IEC approval of 26450 and 26451.

At the European level, CEN/CENELEC JTC25 is developing the EN 18235 series on Trusted Data Transactions: Part 1 (terms and definitions) is in publication as EN 18235-1:2026; Part 2 (trustworthiness requirements) is in its second editorial phase; Part 3 (interoperability requirements) is in preparation for enquiry. A technical specification on maturity assessment of Common European Data Spaces (CEN/CLC/TS 18331) is under vote. The EN 18235 conceptual model defines a transaction lifecycle (grant rights, publication, discovery, negotiation, data exchange, access and usage) that maps directly to the Eclipse DSP protocol flow. The EN 18235 series is explicitly framed for Data Act Article 33 compliance, making these standards commercially and legally significant for organisations.

ETSI is contributing Data Catalog Vocabulary profiles and extensions, with an implementation framework for semantic assets forthcoming. The Institute of Electrical and Electronics Engineers (IEEE) has published IEEE 3800-2024 (Data Trading System: overview, terminology and reference model), the only fully published standard in this landscape. Meanwhile, P3800.1 (protocol and object framework) is in development and 3800.2 (data space discovery protocol) was authorised in February 2026. The IEEE P3800 working group approaches data spaces from a data trading (marketplace-centric) perspective, complementing the provider-control-centric framing of the IDSA and ISO standards. Its chair is the Secretary General of Japan's Data Society Alliance, illustrating how the Japanese institutional ecosystem bridges international standardisation tracks.

2.3. A two-tier working definition

The diversity of definitions described above poses an analytical challenge. A framework broad enough to encompass a too-wide range of data-sharing arrangements as “data spaces”, such as the ISO/IEC DIS 20151 definition, risks losing the specificity needed to assess their distinctive value. Conversely, a narrow definition tied to a particular architecture risks excluding functionally equivalent arrangements that use different terminology or institutional pathways. To navigate this tension, this paper proposes a two-tier working definition.

The **broad definition**, informed by the architecture-agnostic formulations of ISO/IEC DIS 20151 and the DSSC Blueprint, defines a data space as

a data ecosystem in which participants share data through agreed common policies, standards, semantic models and processes, supported by enabling technical and organisational services that facilitate trusted and interoperable data sharing.

This formulation ensures that the OECD’s analytical framework remains inclusive of arrangements that exhibit some functional characteristics of data spaces without adopting the label. Examples include commercial data platforms such as Snowflake and Databricks, which provide cross-organisation data sharing and discovery services, and national statistical systems in which national statistical offices increasingly act as trusted data intermediaries (TDIs) (OECD, 2025^[13]). The global statistical data exchange (SDMX) ecosystem illustrates how a mature data-sharing ecosystem can satisfy the broad definition without originating from the data spaces discussion (Box 2.2).

The **narrow definition** isolates the features whose distinctive value the remainder of this paper assesses. It defines a data space as

a data ecosystem characterised by (i) a federated or decentralised governance framework; (ii) open, standards-based interoperability services; and (iii) mutual trust and effective control between data providers and data users over how data are used.

Both definitions use the term “data ecosystem” as defined in the OECD Recommendation on Enhancing Access to and Sharing of Data [OECD/LEGAL/0463] (OECD, n.d.^[14]), i.e. “the integration of and interaction between different relevant stakeholders, according to their different roles, responsibilities and rights, technologies and business models”. “Federated or decentralised” describes settings in which governance authority and data governance supporting services are not concentrated within a single organisation, without implying that all functions must be fully decentralised. “Open, standards-based interoperability” refers to shared technical services (e.g. identity management, data catalogues, semantic models) that enable participants to discover and exchange data without proprietary lock-in.

Data provider control extends beyond traditional access control (determining who may obtain data) to encompass post-access usage control, i.e. provider-defined conditions on purpose, duration, geographic scope and onward sharing (Reimsbach-Kounatze, 2021^[15]). Their implementations exhibit a gradient of mechanisms – from contractual and purpose-bound restrictions to technically enforced usage constraints. In practice, effective provider control also implies a federated data architecture where data remain with the provider. They are accessed through interfaces such as connectors or other protocol-level endpoints rather than pooled into a central data repository.

Unless otherwise stated, “data spaces” in the remainder of this paper refers to the narrow definition. Figure 2.1 illustrates these features in a stylised architecture of a data space.⁵ The figure is functional rather than prescriptive. In particular, it leaves open where computation occurs,⁶ an architectural choice with direct implications for usage control enforceability and for the integration of complementary tools such as PETs discussed in Section 3.3.3 of Chapter 3.

Box 2.2. SDMX: A mature data-sharing ecosystem for official statistics

Statistical Data and Metadata eXchange (SDMX, ISO 17369), sponsored by seven international organisations (BIS, ECB, Eurostat, IMF, OECD, United Nations, World Bank), has been operational for over two decades. It provides an information model, a standardised REST API, multiple exchange formats and content guidelines that harmonise the meaning of statistical concepts across jurisdictions. Hundreds of national statistical offices and international organisations use SDMX routinely to exchange macroeconomic, social and environmental statistics.

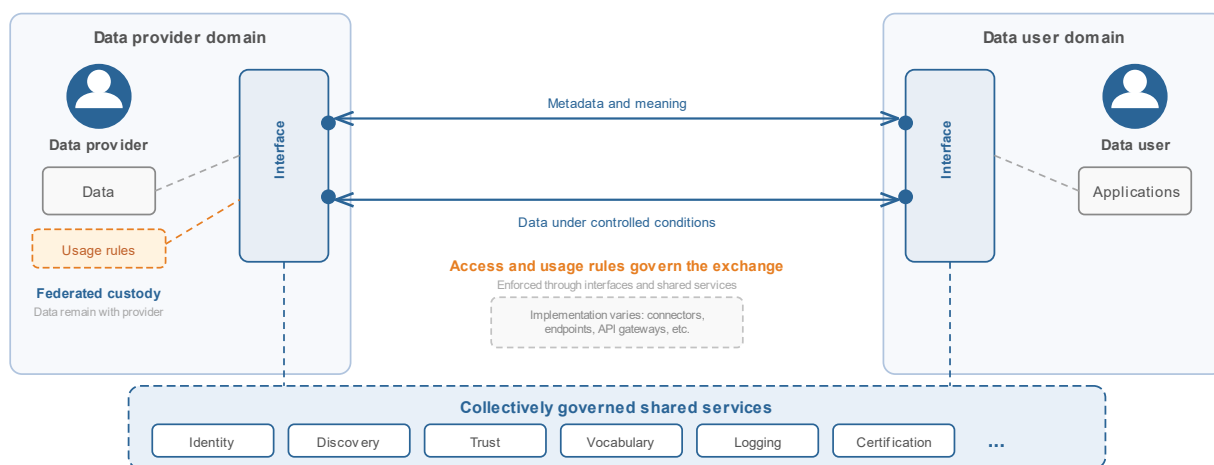
SDMX satisfies each element of the broad definition: agreed common policies (reporting obligations under international frameworks), standards (the SDMX Information Model and API, formalised as an ISO standard), semantic models (globally shared codelists and concept schemes), and processes (regular transmission cycles governed by release calendars). These are supported by enabling services such as the open-source .Stat Suite platform maintained by the Statistical Information System Collaboration Community (SIS-CC).

SDMX does not necessarily satisfy the narrow definition, but it does not have to do so. Aggregate official statistics are published openly, without the need for post-access usage control. The statistical community has developed complementary mechanisms where confidentiality is required for microdata. Such mechanisms include secure processing environments and the use of privacy-enhancing technologies through which code travels to the data rather than the reverse, a pattern functionally equivalent to federated data processing approaches explored in data space initiatives.

Note: BIS = Bank for International Settlements; ECB = European Central Bank; IMF = International Monetary Fund; REST API = Representational State Transfer Application Programming Interface.

Source: SDMX (2026[34]), SDMX – Statistical Data and Metadata eXchange, <https://sdmx.org/>.

Figure 2.1. Stylised architecture of a data space



Note: The figure illustrates the three defining features of the narrow definition used in this paper: federated data custody (data remain with the provider); open, standards-based interoperability (e.g. identity, discovery, vocabulary, trust services); and data provider control (e.g. usage rules govern the conditions of exchange). The "Interface" is an intentional abstraction with implementations varying in whether this function is performed by dedicated middleware (e.g. connectors), protocol-level endpoints, API gateways or other mechanisms.

Source: OECD based on IDSA (2024^[10]), *IDS Reference Architecture Model v4.2*, https://github.com/International-Data-Spaces-Association/IDS-RAM_4_0; DSSC (2026^[16]), *Data Spaces Blueprint v3.0*, <https://dssc.eu/space/BVE/357074549/Data+Spaces+Blueprint+v3.0>; IPA (2026^[5]), *Open Data Spaces Reference Architecture Model v2*, <https://open-dataspaces.gitbook.io/ods-docs/ods-ram/v2>.

2.4. Data spaces and related concepts

The two-tier approach clarifies the relationship between data spaces and adjacent concepts. Two patterns are relevant: some concepts share architectural principles with data spaces but lack the governance features that the narrow definition isolates, while others meet the narrow definition in practice without adopting its terminology.

2.4.1. Adjacent concepts

Commercial data platforms such as Snowflake and Databricks provide cross-organisation query, discovery and sharing services across heterogeneous datasets without requiring full semantic alignment. This is closer to the original dataspace concept by Franklin, Halevy and Maier (2005^[11]) (Section 2.1) than to what “data spaces” now denotes. As such, it falls within the broad definition rather than the narrow one.

A related concept that has gained wide traction in data management practice, *data mesh* (Dehghani, 2019^[17]; Dehghani, 2022^[6]), arrives at the same destination from a different starting point. A decentralised sociotechnical approach for managing analytical data at scale, data mesh is organised around four principles: domain-oriented ownership, data-as-a-product, self-serve data infrastructure and federated governance. While data mesh was originally proposed for intra-organisational settings, its principles are increasingly applied across organisational boundaries.

Both data mesh and the dataspaces concept developed by Franklin, Halevy and Maier reject fully integrated architectures in favour of decentralisation, discoverability and progressive integration. However, data mesh adds what the original dataspace concept lacks: an organisational design layer with domain ownership as an accountability structure, data-as-a-product as a delivery contract and self-serve infrastructure as a platform responsibility.

When extended to inter-organisational settings, both concepts largely converge on the same model that commercial data platforms (Snowflake Marketplace, Databricks Delta Sharing) already implement: cross-organisational data products with platform-mediated governance. In the United States, for instance, an ecosystem of machine builders, manufacturers and integrators has pursued a similar extension under the label “data fabric”. This applies data mesh principles across organisational boundaries through open semantic standards but without the usage control layer (Box 2.3).

These adjacent approaches provide important practical benefits that have contributed to their wide adoption. In particular, they enable scalable and flexible data sharing across organisational boundaries without requiring full centralisation or harmonisation of data. By relying on decentralised ownership, shared semantic definitions and standards-based interoperability mechanisms, they can significantly reduce integration costs and accelerate development of cross-organisational data ecosystems. As such, these architectures often represent pragmatic pathways for organisations to realise value from data sharing.

However, neither the original dataspace concept, nor data mesh, nor inter-organisational data fabric approaches provide the data usage control and collectively governed trust infrastructure that characterise data spaces under the narrow definition. The ODS-RAM v2 explicitly positions itself as inheriting data mesh’s paradigm, extending it to inter-organisational scope. It adds the Identity and Usage Control pillar to address the governance complexity arising when domains cross organisational and national boundaries (Tsuda, 2026^[18]).

The difference between data spaces and these adjacent concepts is thus not about decentralisation (these concepts all have that) or cross-organisational scope (commercial platforms and data mesh provide that). Rather, the differences lie in the features isolated by the narrow definition: whether governance is federated rather than platform-concentrated; whether interoperability is open and standards-based rather than proprietary; and whether data usage control remains with the data providers. The data fabric of the Clean Energy Smart Manufacturing Innovation Institute demonstrates that substantial value can be created from

the interoperability layer alone (Box 2.3); data spaces in the narrow definition add governance and usage control layers on top.

Box 2.3. The “data fabric”: Data mesh extended to inter-organisational settings

The Clean Energy Smart Manufacturing Innovation Institute (CESMII), one of 20 Manufacturing USA institutes, was established in 2016 with USD 70 million by the US Department of Energy and an equivalent amount by the private sector. Bringing together over 200 member organisations, including machine builders, manufacturers and integrators, CESMII pursues a “data fabric” for manufacturing – an approach functionally equivalent to data mesh extended across organisational boundaries. Members contribute standardised data type definitions (Smart Manufacturing Profiles) that function as reusable, vendor-independent data-sharing agreements, which are supported by an open API specification (i3X) for data access.

A joint demonstrator with Germany’s “Platform Industrie 4.0” initiative has validated cross-border interoperability between SM Profiles and the European Asset Administration Shell for a carbon reporting use case. CESMII participates in the International Manufacturing-X (IM-X) Council, which brings together nine founding member organisations from Austria, Canada, France, Germany, Italy, Japan, Korea, the Netherlands and the United States. Spain and Australia participate as additional members.

SM Profiles specify what data mean and how they can be accessed, but not what may be done with them. Data control is achieved through semantic standardisation: the provider controls the model through which data are exposed, rather than through technically enforced data-use constraints. Participation is driven by the value proposition of reduced integration costs. Since 2020, original equipment manufacturers, system integrators and manufacturers have crowdsourced over 100 companion specifications. These cover equipment types across sectors including energy, as well as automotive and semiconductor manufacturing.

2.4.2. Equivalent initiatives

Parallel developments outside the international data spaces ecosystem have independently arrived at arrangements that share the defining features of the narrow definition without adopting its terminology. This was confirmed by participants at the OECD Expert Workshop, who observed that in the United States, for instance, intra-organisational data ecosystems appear to exist in business but are not named as data spaces. This development suggests the concept may need to be framed differently outside the European context where the terminology originated (Annex A).

The Trusted Exchange Framework and Common Agreement (TEFCA) operates as a federated “network of networks” for nationwide health information exchange in the United States, with approximately 500 million records exchanged (HHS, 2026^[19]). Governance is collectively defined under a common agreement independent of any single platform operator, and qualified health information networks enforce usage conditions set by the data provider. In India, Data Empowerment and Protection Architecture (DEPA) provides a consent-based federated data-sharing framework built on India Stack. This is operational in financial services through the Account Aggregator framework, which intermediates consent-driven data flows across more than 270 million linked accounts; usage control is enforced structurally through the consent architecture.⁷

The iSHARE Trust Framework, originating in the Dutch logistics sector and used across multiple sectors and EU jurisdictions (iSHARE Foundation, n.d.^[20]), is a further example. iSHARE provides a scheme-based approach to identity, authentication and authorisation for data sharing. As such, it uses operator certification rather than a mandated connector implementation, and a community-governed wiki model to

develop shared use cases. iSHARE illustrates an alternative architectural emphasis: governance and trust scheme membership as the primary co-ordination mechanism. Technical implementation choices are left more open to participants than in connector-centric models such as IDS.

TEFCA and DEPA were not designed as data spaces, yet both exhibit the defining features of the narrow definition: federated governance independent of any single platform operator; open and standards-based interoperability; and structurally enforced usage control. Their existence confirms the inclusive scope of the broad definition is warranted, and that the narrow definition identifies structural properties rather than a label (Section 4.1 of Chapter 4 examines the implications of this convergence for the co-ordination problem).

References

- Dehghani, Z. (2022), *Data Mesh: Delivering Data-Driven Value at Scale*, O'Reilly. [6]
- Dehghani, Z. (2019), "How to move beyond a monolithic data lake to a distributed data mesh", [17]
<https://martinfowler.com/articles/data-monolith-to-mesh.html> (accessed on 15 April 2026).
- DSSC (2026), "Data Spaces Blueprint v3.0", [16]
<https://dssc.eu/space/BVE/357074549/Data+Spaces+Blueprint+v3.0> (accessed on 9 April 2026).
- ETSI (2025), "Draft ETSI EN 304 199 V1.0.0 (2026-03): Data Solutions (DATA); Data catalogue implementation framework; Guidelines for Data Catalogue Framework", ETSI. [3]
- Franklin, M., A. Halevy and D. Maier (2005), "From databases to dataspace", *ACM SIGMOD Record*, Vol. 34/4, pp. 27-33, <https://doi.org/10.1145/1107499.1107502>. [1]
- Halevy, A., M. Franklin and D. Maier (2006), "Principles of dataspace systems", *Proceedings of the twenty-fifth ACM SIGMOD-SIGACT-SIGART symposium on Principles of database systems*, pp. 1-9, <https://doi.org/10.1145/1142351.1142352>. [23]
- HHS (2026), "TEFCA – America's national interoperability network reaches nearly 500 million health records exchanged", <https://www.hhs.gov/press-room/tefca-americas-national-interoperability-network-reaches-nearly-500-million-health-records-exchanged.html> (accessed on 7 April 2026). [19]
- IDSA (2024), *IDS Reference Architecture Model, Version 4.0*, https://github.com/International-Data-Spaces-Association/IDS-RAM_4_0. [10]
- IOFDS (n.d.), "International Open Forum on Data Society", <https://iofds.org/> (accessed on 13 April 2026). [11]
- IPA (2026), "Open Data Spaces Reference Architecture Model (ODS-RAM)", <https://open-dataspaces.gitbook.io/ods-docs/ods-ram/v2> (accessed on 1 April 2026). [5]
- IPA and Catena-X (2025), "Catena-X and Ouranos ecosystem successfully demonstrate data space interoperability", <https://catena-x.net/news/catena-x-and-ouranos-ecosystem-successfully-demonstrate-data-space-interoperability/> (accessed on 6 March 2026). [12]

- iSHARE Foundation (n.d.), “Foundation History”, <https://ishare.eu/about/the-foundation/foundation-history/> (accessed on 28 April 2026). [20]
- ISO (2025), “ISO/IEC DIS 20151-1: Cloud computing and distributed platforms – Dataspaces – Part 1: Concepts and characteristics”, <https://www.iso.org/standard/86589.html> (accessed on 14 April 2026). [9]
- Kernstock, P. et al. (2025), “Establishing and governing data ecosystems at the crossroads of centralization and decentralization”, *Electronic Markets*, Vol. 35/71, <https://doi.org/10.1007/s12525-025-00810-x>. [22]
- METI/IPA (2025), *Open Data Spaces Reference Architecture Model (ODS-RAM)*, Ministry of Economy, Trade and Industry (METI) and the Information-Technology Promotion Agency, Government of Japan. [4]
- NSF (2023), *NSF Efforts to Achieve the Nation’s Vision for the Materials Genome Initiative: Designing Materials to Revolutionize and Engineer Our Future (DMREF)*, National Academies Press, Washington, DC, <https://doi.org/10.17226/26723>. [8]
- OECD (2025), “National statistical offices as emerging trusted intermediaries in data governance”, *OECD Digital Economy Papers*, No. 378, OECD Publishing, Paris, <https://doi.org/10.1787/3721ec38-en>. [13]
- OECD (n.d.), *Recommendation of the Council on Enhancing Access to and Sharing of Data*, OECD/LEGAL/0463, <https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0463> (accessed on 6 March 2023). [14]
- Otto, B. et al. (2017), *Reference Architecture Model for the Industrial Data Space*, Fraunhofer-Gesellschaft, Germany, <https://doi.org/10.24406/PUBLICA-FHG-298818>. [2]
- Peeters, S. (2013), “Beyond distributed and decentralized: What is a federated network?”, <https://networkcultures.org/unlikeus/resources/articles/what-is-a-federated-network/> (accessed on 22 April 2026). [21]
- Reimsbach-Kounatze, C. (2021), “Enhancing access to and sharing of data: Striking the balance between openness and control over data”, *Data Access, Consumer Interests and Public Welfare*, pp. 25-68, <https://doi.org/10.5771/9783748924999-25>. [15]
- Steinbuss, S. et al. (2021), “Usage control in the international data spaces”, *Position Paper*, International Data Spaces Association, Dortmund, Germany, <https://doi.org/10.5281/zenodo.5675884>. [7]
- Tsuda, M. (2026), *Why Open Dataspaces: Design Philosophy and Architectural Paradigm*, Information-technology Promotion Agency, Tokyo, https://www.ipa.go.jp/en/digital/architecture-guidelines/individual-link/p1o1lf000001xv4n-att/WhyOpenDataspaces_en.pdf. [18]

Endnotes

¹ Crucially, partial integration is treated as a legitimate resting state, not merely a waypoint towards full alignment; each participant invests in further integration only where the marginal return justifies the marginal cost. The subsequent formalisation of Dataspace Support Platforms (DSSPs) by Halevy, Franklin and Maier (2006^[23]) sharpened the distinction: whereas traditional data integration systems require complete semantic mappings as a precondition, dataspace emphasise coexistence, incrementalism and best-effort functionality.

² Much of the data spaces literature uses “decentralised” to describe this architecture. This paper uses “federated” as the more precise term: participants retain autonomy over their data but rely on collectively governed shared services (identity, discovery, transaction logging) for co-ordination (Peeters, 2013^[21]). The IDS RAM v4.0 itself describes “a federated structure, in contrast to completely centralized or decentralized architectures” (IDSA, 2024^[10]); Kernstock et al. (2025^[22]) document empirically how governance functions are distributed along a gradient rather than located at either pole. The distinction matters analytically: federated architectures require collective investment in shared services, which is the source of the bootstrapping challenges discussed in Chapter 3.

³ The first version of Japan’s Open Data Spaces Reference Architecture Model adopted this definition verbatim in footnote 3 of METI and IPA (2025^[4]); ODS-RAM v2 uses a different formulation.

⁴ The concept of data sovereignty is thus more specific and distinct from the geopolitical concept of sovereign jurisdiction over data flows. This paper uses the term “data provider control” to avoid conflation with the geopolitical usage.

⁵ The working definition and Figure 2.1 as adopted in this report is functional rather than architectural, with the aim to ground the policy analysis that follows. The working definition and Figure 2.1 are therefore limited to the scope of the report, which concerns the use of data spaces as typically foreseen in data sharing scenarios between entities, where the core functional elements are invariant: a data provider; a data user; a data transfer mechanism; a set of services that enable trust, interoperability and discoverability, and more; and this with the main purpose of discussing the viability and adoption challenges of data spaces.

⁶ A further architectural choice, not specific to data spaces but increasingly relevant within them, is the locus of computation. Three patterns coexist: (i) *data-to-code*, where data are transferred to the user’s environment for local processing; (ii) *code-to-data*, where queries or model-training rounds travel to the provider and only results return (e.g. federated learning); and (iii) computation in a *neutral compute environment* such as a trusted execution environment or data clean room operated as a shared service. The choice has direct implications for usage control enforceability (Section 3.3.3 of Chapter 3), the integration of privacy-enhancing technologies, and the distribution of compute costs among participants.

⁷ DEPA’s locus of control differs from the narrow definition in two respects. First, usage conditions are defined by the data subject (the individual) rather than the data provider (the financial institution). Second, governance is regulator-defined (RBI licensing) rather than collectively negotiated among participants. DEPA thus shares the architectural pattern (federation, structurally enforced usage control, no centralised data repository) while serving a different governance logic: data subject control, i.e. individual data empowerment, rather than data provider control. [See DSTI/DPC/DGP(2025)8/REV1.]

3

Multi-party co-ordination as the distinctive value proposition and challenge of data spaces

This chapter identifies the distinctive value proposition of data spaces. It characterises multi-party data sharing as a collective-action problem that neither market pricing nor centralised co-ordination resolves, and defines the boundary conditions under which data spaces hold a comparative advantage over bilateral contracts, centralised platforms and privacy-enhancing technologies. It examines the three mechanisms through which data spaces address this problem: distributed governance, open interoperability and provider-defined usage control. The chapter concludes with the bootstrapping problem that explains much of the gap between investment and operational viability, drawing on case studies of successful and failed initiatives.

Data spaces target a specific class of data-sharing problem, namely multi-party co-ordination in sectors where no single participant holds all the data needed for system-level outcomes. Section 3.1 discusses this co-ordination problem behind multi-party data sharing. Section 3.2 locates data spaces relative to alternative mechanisms – such as bilateral contracts, centralised platforms and services based on privacy-enhancing technologies (PETs) – and defines the boundary conditions under which data spaces have a comparative advantage. Section 3.3 examines the three operational mechanisms through which data spaces address the co-ordination problem: (i) distributed governance; (ii) open interoperability; and (iii) provider-defined usage control. Section 3.4 turns to a second-order challenge that explains the maturity and adoption gap to a significant extent.

3.1. The co-ordination problem behind multi-party data sharing

Multi-party data sharing is, at its root, a collective-action problem (Olson, 1965^[1]). Multiple parties must co-ordinate to produce a shared good despite individual incentives that work against collective provision. The two mechanisms that ordinarily co-ordinate economic activity – market pricing and centralised authority (Williamson, 1985^[2]; Williamson, 1991^[3]) – both face structural obstacles in this context.

The failure of price formation: in standard markets, prices co-ordinate allocation without requiring institutional intervention. However, data shared within multi-party ecosystems cannot be efficiently priced because their values are inherently context-dependent.¹ This makes standardised valuation across participants structurally difficult (see Box 1.1 of Chapter 1).² A component traceability record in an automotive supply chain, for instance, has different value to the tier-1 supplier, the original equipment manufacturer (OEM), the recycler and the regulator. The total social value exceeds the sum of bilateral values because all parties having access enables system-level outcomes (supply chain resilience, regulatory compliance, circular economy integration) that no single data transaction captures. Frischmann (2012^[4]) argues that resources exhibiting these properties generate social value that price mechanisms alone may not fully capture. This does not imply that market pricing for data is generally unworkable: commercial data markets and data-brokerage services demonstrate that price-mediated exchange functions where data products are sufficiently standardised. The claim here is narrower: in multi-party ecosystems where value is collective, context-dependent and not assignable to a single rights-holder, the conditions for efficient price formation are weakened.

The failure of centralised co-ordination: where prices cannot help co-ordinate, a vertically integrated or dominant actor typically steps in, setting the terms and internalising the co-ordination costs. In multi-party data sharing, this route tends to be blocked wherever the participants are competitors or operating in different jurisdictions and refuse to cede governance authority to a rival-adjacent or otherwise untrusted data intermediary. The most instructive historical case is Covisint, an automotive data exchange platform. Launched in 2000 by General Motors, Ford and DaimlerChrysler (and later joined by Renault and Nissan), Covisint was designed to centralise procurement and supply chain data across the industry. Suppliers resisted, seeing the platform as a tool for buyer leverage over pricing and inventory rather than as neutral infrastructure (Arbin and Essler, 2005^[5]; Gerst and Bunduchi, 2005^[6]). The auction business was sold off, the founding OEMs withdrew and the original exchange model was abandoned by 2004. The failure was not technological but rather institutional. Various waves of supply chain data-sharing approaches – from electronic data interchange to integration based on service-oriented architecture and, more recently, centralised control towers – each encountered similar governance limitations at a new level of abstraction.³

The compounding of both failures by asymmetric cost-value distribution among participants: in multi-party data sharing, those who bear the costs of contributing data frequently do not capture proportional value from their use (Gelhaar et al., 2021^[7]; Brousseau, Toledano and Eustache, 2024^[8]). Tier-2 automotive suppliers, for instance, invest in data collection and formatting for traceability records whose primary value accrues to OEMs and regulators; farmers bear the digitisation costs of generating

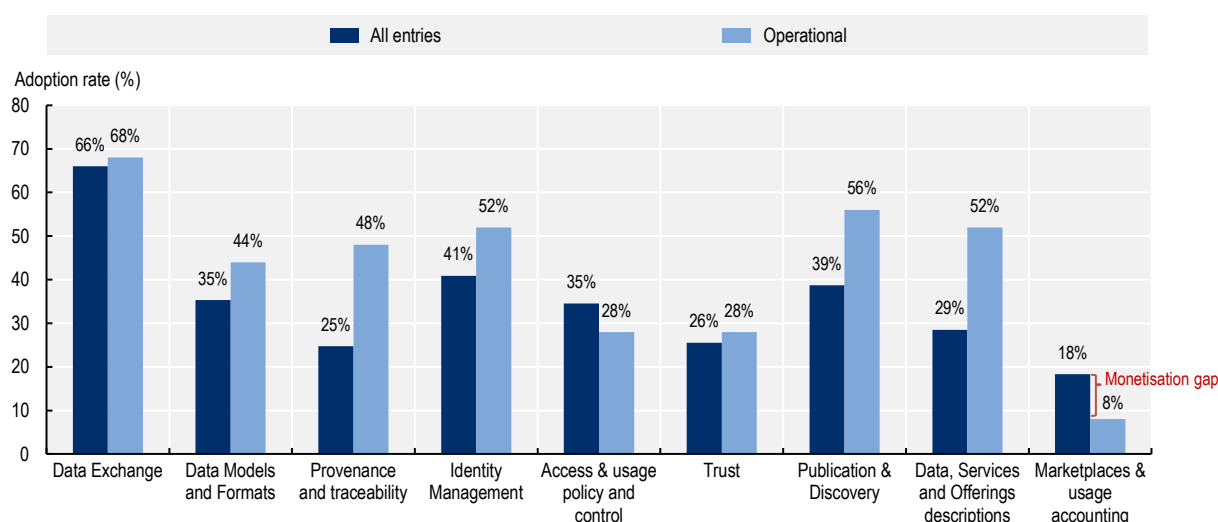
field-level data while downstream processors and technology providers capture most of the analytical value. Practitioners at the OECD Expert Workshop confirmed this imbalance constitutes a persistent challenge (Annex A). The asymmetry is not merely distributional; it is a structural participation barrier that directly affects the second level of the bootstrapping problem analysed in Section 3.4. This is because data providers whose expected returns fall below their contribution costs will rationally under-provide data regardless of the governance model.

For government ministries and public administrations, the absence of price-mediated co-ordination is the default condition rather than a failure mode. Public-sector data sharing does not operate on a profit-and-loss logic. Rather, the return on investment of a public sector data space is measured in public value generation, such as reduced administrative costs, cross-agency policy coherence and improved delivery of services to citizens (OECD, 2019^[9]). This is the limiting case of Frischmann's (2012^[4]) argument above. Where infrastructural resources in commercial settings generate social value that price mechanisms capture only partially, the value generated by public-sector data sharing can almost entirely consist of non-market goods and therefore escapes price formation altogether. The co-ordination problem can nonetheless persist in this setting. It typically takes the form of institutional fragmentation across ministries with separate mandates, budgets and accountability structures, rather than as failed price formation.⁴ Section 4.5 and Box 4.2 of Chapter 4 discuss public sector data-sharing frameworks and their public value logic in more detail.

Data space architectures decompose the capabilities needed for multi-party data sharing⁵ into modular functional components, commonly referred to as “building blocks”. These include components that correspond to the defining features of the narrow definition (Section 2.3 of Chapter 2), such as data exchange infrastructure and usage control. They also include additional capabilities such as publication, discovery and marketplace functions that support the commercial layer through which participants derive value from shared data. The DSSC Blueprint (DSSC, 2026^[10]) provides the most widely adopted taxonomy, grouping building blocks into three categories: data interoperability, data sovereignty and trust, and data value creation. The International Data Spaces Association (IDSA) Radar tracks self-reported adoption of these building blocks across initiatives, providing an empirical proxy for which capabilities data spaces prioritise at different maturity stages (Figure 3.1).

The empirical evidence is consistent with these types of failures. The marketplace building block, which embodies price-mediated data exchange, declines from 18% adoption across all IDSA Radar initiatives to 8% at operational stage. This suggests that operational data spaces tend to move away from marketplace logic as they mature (Figure 3.1). As the sample of operational data spaces is small (25 entries), and building block adoption is self-reported (Box 1.1 of Chapter 1), the pattern should be read as indicative rather than conclusive. It is, however, consistent with case study analysis. The data spaces that have achieved higher maturity, such as the Finnish Maritime Data Space (MDS), operate on cost sharing and reciprocal exchange models under neutral governance rather than on data monetisation through a centralised intermediary. They treat shared data as infrastructure to be jointly used and maintained rather than as a commodity to be priced or controlled by a single actor.

Figure 3.1. Building block adoption rates: All entries vs. operational data spaces



Note: Adoption rates are computed by checking whether each building block name appears in the concatenated multi-value string reported by each entry. The denominator is the full population (235 entries for “all” and 25 for “operational”). Building blocks follow the categories defined in the DSSC Blueprint: data interoperability (data exchange, data models and formats, provenance and traceability), data sovereignty and trust (identity management, access and usage policy and control, trust), and data value creation (publication and discovery, data services and offerings descriptions, marketplaces and usage accounting).

Source: Based on IDSA (n.d.), *IDSA Data Spaces Radar*, www.dataspaces-radar.org.

3.2. Comparative advantage

When market pricing cannot co-ordinate (data) allocation and no single actor is trusted, the co-ordination problem identified in Section 3.1 requires an organisational form that is neither purely market-based nor centralised. Williamson (1985^[2]; 1991^[3]) argues precisely that for transactions characterised by high asset specificity, uncertainty and frequency (all three of which characterise multi-party data exchange via data spaces), “hybrid” governance forms emerge as a distinct organisational response. This “is not a loose amalgam of market and hierarchy but possesses its own disciplined rationale” (Williamson, 1991^[3]).⁶ Data spaces provide one such hybrid governance form (as discussed in Section 3.3.1).

Four specific boundary conditions define the class of problem (use cases) for which data spaces have a comparative advantage over the alternative mechanisms compared in Table 3.1 (e.g. bilateral contracts, centralised platforms and approaches based on PETs). These conditions are (i) three or more parties; (ii) competitive or jurisdictional relationships among them that preclude trusting a single commercial intermediary; (iii) non-excludable collective value that bilateral contracts cannot capture; and (iv) ongoing data flows requiring continuous governance rather than one-off transactions. Where these conditions are absent, simpler data-sharing mechanisms may be sufficient and may not justify the co-ordination costs that data spaces entail.

Table 3.1. Comparison of data-sharing mechanisms and enablers

Dimension	Bilateral contracts	Centralised platforms	Data-sharing services based on PETs	Data spaces (narrow definition)
Governance distribution	Pairwise; no shared governance layer	Concentrated in platform operator	Typically platform-operated or bilateral	Distributed among participants; neutral intermediary (non-profit, state enterprise, consortium or commercial operator)
Interoperability	Bespoke point-to-point; rebuilt per partner	Proprietary connectors and formats; platform lock-in	Vendor-specific enclaves; limited cross-platform	Standards-based protocols and open data models (e.g. Eclipse Dataspace Protocol and Open Data Spaces Protocol)
Portability	High (no infrastructure dependency)	Low (switching costs increase with usage)	Low (computation tied to PET provider)	High (standards-based; investment in data products and policies is portable)
Access control	Legal identification clauses; pairwise	Platform IAM; platform-mediated	Secure enclaves; participant authentication	Decentralised or federated identity (e.g. verifiable credentials in IDS; OpenID Connect in ODS)
Usage control	Legal clauses; no technical enforcement; does not scale	Platform ToS govern post-access use; provider cedes control	Structural (data never leave enclave); limited to computation-only	Provider-defined conditions on post-access use; implementations combine contractual commitments with machine-readable policies (e.g. Open Digital Rights Language) and, for stronger enforcement, PETs or data sandbox environments
Regulatory alignment	Neutral (depends on contract terms)	Structural tension with EU Data Governance Act neutrality requirements; platform ToS may conflict with EU GDPR purpose limitation	Partial (supports confidentiality and privacy regulation; but may not address governance requirements)	Structurally aligned with regulatory frameworks requiring multi-party traceability, interoperability and provider control (e.g. DPPs)
Scalability	Combinatorial: $n(n-1)/2$ agreements for n parties	High within platform boundary; low across platforms	Moderate (e.g. multi-party computation can be costly)	Moderate (co-ordination costs; collective-action challenges in bootstrapping)
Co-ordination cost	Low per agreement; high in aggregate	Low (platform absorbs)	Low (technical setup)	High (governance establishment, standard adoption, trust building)
Value capture model	Price negotiation per transaction; captures only bilateral value	Platform-mediated pricing; cross-subsidisation across sides; captures monetisable value	Computation fees; value of insight, not underlying data	Cost sharing, reciprocal exchange or public subsidy; captures collective and infrastructural value that bilateral pricing cannot capture

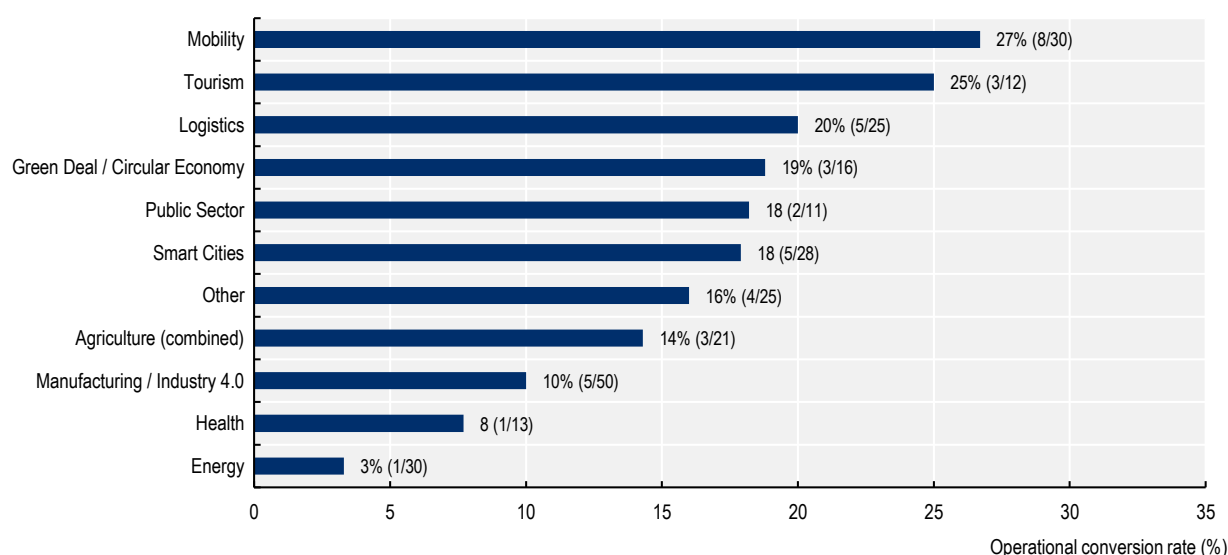
Note: IAM = Identity and Access Management Platform; DPP = Digital Product Passport; EU GDPR = General Data Protection Regulation; IDS = Industrial Data Space; ODS = Open Data Spaces; ToS = Terms of Service. The table compares idealised properties of each mechanism along dimensions relevant to the boundary conditions identified in Section 3.2. The data spaces column reflects general properties of the narrow definition; specific implementations vary across architectural approaches (e.g. IDS-RAM/EDC, ODS-RAM). Where illustrative examples are given, they are not exhaustive.

The alternative mechanisms in Table 3.1 each serve legitimate purposes where one or more of these boundary conditions is absent, although none addresses all four simultaneously.

- Bilateral contracts are the default mechanism for pairwise data sharing, but their co-ordination costs scale combinatorially with the number of participants [$n(n-1)/2$ for n parties], making them impractical where boundary condition (i) applies. They also provide no shared governance layer through which collective value, boundary condition (iii), can be created or allocated.
- Centralised commercial platforms have built successful multi-party data exchange businesses. Snowflake, AWS Data Exchange and Databricks together serve tens of thousands of customers; roughly a quarter of Snowflake's customers maintain sustained cross-account data-sharing relationships. Many organisations evidently accept vendor terms of service in exchange for functionality, speed and network effects. These platforms succeed by offering a different trade-off, forgoing data provider control and multi-stakeholder governance in exchange for lower co-ordination costs and immediate functionality.
- PETs address the confidentiality dimension of multi-party data sharing, enabling computation on combined data without revealing individual inputs. They do not, however, address the institutional co-ordination problem of who governs, who pays and how collective value is allocated (see Section 3.3.3 for how PETs complement rather than substitute for data space governance).

Grouping 147 IDSA Radar initiatives by their self-reported value propositions yields a pattern consistent with the boundary conditions identified above. The five domains where data spaces have a partially unique structural advantage (supply chain visibility across competitors, health data exchange across jurisdictions, multimodal mobility co-ordination across operators, energy system integration across prosumers and utilities, and smart city data sharing across public and private providers) all share the structural feature identified above: they are n -party co-ordination problems where no commercial participant is trusted to serve as intermediary. Figure 3.2 provides a quantitative complement, showing the operational share by sector. The sectors with the highest operational share tend to align with the domain clusters identified above.

Figure 3.2. Operational share by sector



Note: "Operational conversion rate" is the number of IDSA Radar entries that have reached operational status as a proportion of all entries in that sector. Sectors are self-reported and multi-valued: a single entry may appear in more than one sector, so sector totals exceed 235. "Agriculture" combines entries labelled "Agriculture/Agrifood" (14) and "Agriculture" (7) in the Radar. Only sectors with three or more total entries are shown.

Source: Based on IDSA (n.d.), *IDSA Data Spaces Radar*, www.dataspaces-radar.org.

3.3. How data spaces address the co-ordination problem of multi-party data sharing

Data spaces address the multi-party co-ordination problem through three operational mechanisms. These operational mechanisms are in line with the narrow definition (Section 2.3 of Chapter 2) and their ordering below reflects how adoption proceeds in practice, informed by case study evidence:

- Distributed governance (Section 3.3.1) provides the institutional foundation.
- Open interoperability (Section 3.3.2) provides the technical infrastructure that participants actually deploy.
- Provider-defined usage control embedded as a capability in the interoperability stack. The trust capability (Section 3.3.3), comprising in particular identity and data usage control, distinguishes data space interoperability from both commercial platform interoperability and traditional industry consortium data exchange.

The distinctive feature is not any single mechanism but their integration across three facets commonly distinguished in the dataspace literature (i.e. technical, business and legal). Through this integration, governance rules become technically enforceable through the interoperability stack; contractual usage conditions become machine-readable and portable across the ecosystem; and business arrangements are anchored to a shared institutional framework. No alternative data-sharing mechanism integrates institutional, technical and contractual controls into a single operational framework as tightly as data spaces.

Participants at the Expert Workshop highlighted the potential that data spaces offer to address challenges in cross-border use cases, amid increased risks of barriers to cross-border data flows (Annex A). In particular, they highlighted that data space architecture allowed participants to negotiate machine-readable policies (at the level of data usage rights). These can be configured to be compatible with the regulations and laws of multiple jurisdictions simultaneously, making these conditions available to all ecosystem participants through the standardised protocol. Where bilateral contracts would require renegotiating terms for each regulatory combination and each two parties, data spaces can in principle encode cross-jurisdictionally compatible requirements once and propagate them across the ecosystem. In so doing, data spaces significantly reduce transaction costs.

IDSA Radar data underscore the difficulty of activating all three mechanisms simultaneously. Of the 153 initiatives at implementation and operational stages, only 45 score on all three operational mechanisms (distributed governance, open interoperability and usage control) (see Figure 3.1). The rarity of multi-mechanism deployment is consistent with the argument that each mechanism addresses a distinct dimension of the co-ordination problem, and partial activation limits the value a data space can deliver.⁷

3.3.1. Distributed governance

Despite the significant importance of distributed governance, in every mature data space examined for this paper, a neutral intermediary or orchestrator occupies a key governance role: a state enterprise (Fintraffic), a non-profit association (Catena-X Automotive Network e.V., see Box 3.1) or an industry consortium (BoostAeroSpace). The pattern is aligned with Guggenberger, Schlueter Langdon and Otto (2025^[11]), who characterise data spaces as “meta-organisations” in which legally independent entities voluntarily co-ordinate under a shared framework. Concentrating governance in a single commercial intermediary would convert the arrangement into a hierarchy, potentially destroying the voluntary co-ordination that defines it. Even data ecosystems that start on decentralised principles tend to establish central institutions as they scale (Kernstock et al., 2025^[12]), but these institutions must remain structurally neutral.

The drift towards centralisation reflects well-documented co-ordination failures: as the network grows, the costs of distributed decision making (rule-setting, dispute resolution, onboarding) rise faster than its

benefits. Consequently, individual participants face incentives to free-ride on shared infrastructure rather than contribute to it. A neutral central institution lowers these transaction costs and concentrates accountability, but at the price of reintroducing the very governance asymmetry that decentralised principles were designed to avoid. This structure enables competitors to share data where no single participant is trusted as data intermediary, or accommodates public-interest data flows.

The neutral intermediary or orchestrator also need to understand both the overall value proposition and the addressable market of the data space in order to initiate and sustain collective action. In particular it needs to address how each stakeholder tends to only understand its own business model in relation to the data space while missing a view of the overarching business value. However, distributing the benefits created by a data space requires complex negotiation among members, including orchestrators who may also have their own business interests (Annex A). The governance challenge is therefore not simply to ensure neutrality in the abstract, but rather to design institutional arrangements that harness stakeholders' incentives without allowing them to capture disproportionate value. The neutrality requirement of the EU Data Governance Act (DGA) can be understood, in part, as a regulatory response to this tension, although it is under revision (see Section 4.1 of Chapter 4).

Where price formation for data fails (Section 3.1), distributed governance can provide an alternative resource allocation mechanism. In this context, rather than pricing individual data transactions, participants collectively negotiate how to share the costs of maintaining shared infrastructure and how to distribute the value it generates. The Finnish MDS illustrates this most directly: its value proposition is cost sharing (quantifiable reductions in fuel costs, port fees and emissions penalties) rather than data monetisation. Lessons learnt from BoostAeroSpace show a similar logic through membership-based cost allocation. Both treat shared data as infrastructure to be collectively maintained, consistent with Frischmann (2012^[4]).

Box 3.1. The case of the Catena-X Automotive Network

Catena-X Automotive Network originated from industry initiatives in 2020-2021. It was established as the association Catena-X Automotive Network e.V. on 15 May 2021 by 17 German founding members, including ARENA2036 and the Fraunhofer Institutes and Research Units (*Fraunhofer-Gesellschaft*) alongside 15 companies (Catena-X, 2023^[13]). As of November 2024, there were 193 members. Catena-X as a data ecosystem has three parts. The first is the Catena-X Automotive Network e.V., which is responsible for standards, certifications and governance. Members contribute to developing the data ecosystem, including its standards, and are obliged to comply with the latter. The second part is the development environment, which proposes candidate standards and develops open-source reference implementations. Finally, the operating environment offers various open-source and commercial services and applications.

The values offered by Catena-X are reflected in its “10 Golden Rules” (Catena-X, 2025^[14]), which are highlighted as contributing to the bigger pictures of European data space strategies: (i) values of Gaia-X and IDSA; (ii) data sovereignty; (iii) governance framework; (iv) freedom of contract; (v) registration and certification; (vi) compliance by conduct; (vii) compliance by design; (viii) security; (ix) control; and (x) legal provisions and disclaimer. In particular, Catena-X is presented as an effective means to ensure compliance with data- and sector-specific regulations, improve supply chain management, and enable secure and efficient data sharing. These values are expected to be further enhanced as more participants join Catena-X through network effects.

Industry-owned consortia such as SWIFT⁸, SITA⁹ and GS1¹⁰ have co-ordinated multi-party data exchange across competing firms for decades without the data space architecture. What data spaces add is a technical enforcement layer (including provider-defined, machine-readable usage control that travels with data) and provenance infrastructure that makes such control verifiable (Section 3.3.3). BoostAeroSpace

illustrates the trajectory: established in 2011 as a governance consortium for aerospace supply chain data exchange, it subsequently adopted data space components (Eclipse Dataspace Components, IDSA architecture) to add technical enforcement to its institutional framework.

An illustrative case is EONA-X, an operational data space in the mobility sector that brings direct competitors under a neutral data intermediary for sustainability data. The Finnish MDS, led by Fintraffic (a state-owned enterprise with an indefinite time horizon), demonstrates business-model-first design: technology choices follow from business requirements.

The IDSA Radar data reflect a related pattern: EU funding accounts for about 46% of implementation-stage entries but drops to about 20% at operational stage (Box 1.1 of Chapter 1). Two interpretations are consistent with this pattern: data spaces that reach operational maturity may develop funding models independent of public grants, or EU-funded initiatives may be more likely to fail before reaching operational stage. The cross-sectional data cannot distinguish between these dynamics, and both may operate simultaneously. Either way, the pattern suggests that EU grant funding alone is not sufficient to sustain data spaces through to operational viability.

The case studies above illustrate what these models look like in practice: membership-based cost allocation (BoostAeroSpace), cost sharing tied to measurable operational savings (Finnish Maritime, see Box 3.2 and Annex A),¹¹ or industry association funding (EONA-X). Notably, none of these models relies on data commercialisation. The decline in marketplace building block adoption from 18% across all initiatives to 8% at operational stage (Figure 3.1) suggests that data monetisation through pricing and trading has not proven a viable funding model for operational data spaces. The implications of this funding transition for policy design are discussed in Section 4.2 of Chapter 4.

3.3.2. Open interoperability and portability

If participants must adopt a proprietary technology stack to exchange data, the infrastructure provider risks becoming a de facto central authority. Standards-based protocols significantly reduce this risk of capture and ensure participants retain the ability to exit, switch and connect across domains.

Practitioner accounts at the OECD Expert Workshop suggest that interoperability is the primary reason participants adopt data space components. Organisations deploy data space components to solve a concrete problem: exchanging data across organisational boundaries without rebuilding integrations for each new partner. The governance framework (Section 3.3.1) determines *why* participants join; the interoperability stack determines *what they actually deploy*. Usage control capabilities (Section 3.3.3) come embedded in this stack, available when needed but not the primary adoption motivation in many cases.

Data spaces specify data exchange through standardised protocols that define a common transaction process (with a possible negotiation phase as foreseen in the Eclipse Dataspace Protocol or optionally provided in the Open Data Spaces Protocol). Reference implementations provide interoperable interfaces (e.g. connectors) and standardised information models ensuring semantic interoperability (see Box 2.1 of Chapter 2). Emerging AI-based semantic mediation may, over time, reduce reliance on rigid standardised information models by translating across heterogeneous schemas at runtime. However, the governance and provenance guarantees required in regulated dataspace still favour explicitly agreed models. Open interoperability keeps switching costs low and makes investment in data products portable across vendors. It also enables multi-domain value propositions that proprietary platforms cannot support without bilateral agreements with every other platform.

The Data Space for Traffic Measurements, an operational initiative in the Flanders region, illustrates standardised protocols by unifying more than 500 entities through a linked data standard; interoperability is the primary value mechanism. Catena-X deploys Eclipse Dataspace Components and standardised information models across the entire automotive value chain, making usage control portable across the

supply chain. A policy attached at the tier-2 supplier level travels with data through to the OEM because the underlying protocol is standardised.

Adoption data show consolidation around open standards within data spaces. At the operational stage, about a third use EDC, while the FIWARE connector drops from about 20% overall to none at operational stage. This pattern is consistent with consolidation around the protocol stack that supports the full Eclipse Dataspace Protocol negotiation and transfer state machine. However, it may also reflect a reporting bias: IDSA maintains the IDSA Radar, and initiatives using alternative connector technologies may be less likely to register (see Box 1.1 of Chapter 1).

The pattern also illustrates the role of open-source implementations as a de facto standardisation mechanism in this space. The Eclipse Dataspace Components and FIWARE codebases are themselves the reference points around which interoperability converges. This lowers the cost of multi-vendor adoption relative to purely specification-led standards.

Architectural decoupling is another route to open interoperability with the ODS-RAM a notable example. The architecture of ODS-RAM is organised around three pillars: ontology and semantic interoperability; data addressability and discoverability; and identity and usage control. Each combines technical and rights-based functions as loosely coupled, independently replaceable components (IPA, 2026^[15]; Tsuda, 2026^[16]).¹² This pairing allows the ontology layer to evolve independently of the data layer. In this way, semantic models can be updated without re-engineering data pipelines or breaking interfaces. In addition, the pairing supports incremental, opt-in adoption – “detachability” in Tsuda (2026^[16]). This is directly relevant to the bootstrapping problem discussed in Section 3.4. Emerging approaches extend this principle through “dynamic ontologies”, which use large language models to automate mappings across heterogeneous ontologies.¹³

The practical consequence is that switching costs can be reduced not only across vendors implementing the same protocol, but also across architectural components. An organisation can replace its identity mechanism or semantic model without re-engineering its data transfer pipeline. The IPA-Catena-X proof of concept demonstrated product carbon footprint data exchange between ODS-RAM- and EDC-based implementations for EU Battery Regulation compliance. In so doing, it provides early evidence that protocol- and architecture-level interoperability is achievable (IPA and Catena-X, 2025^[17]). (See Section 4.3 of Chapter 4).

Commercial platforms show signs of moving in a similar direction at the protocol layer. Databricks open-sourced its Delta Sharing protocol and Unity Catalog, suggesting the economic logic of open standards can override the lock-in incentive even for commercial actors. This is not convergence on data space architecture: Delta Sharing does not implement the Eclipse Dataspace Protocol or the Open Data Spaces Protocol. Moreover, open-sourcing a protocol does not entail adopting distributed governance or provider-defined usage control. It does, however, narrow the gap on the interoperability dimension. This means the structural differentiation of data spaces increasingly rests on governance neutrality (Section 3.3.1) and the trust capability (Section 3.3.3). It is the integration of these dimensions with the open interoperability layer that constitutes the comparative advantage. Governance neutrality is available through public bodies, usage control through bilateral contracts, and open standards through industry consortia. However, no alternative mechanism combines all three.

3.3.3. Identity, usage control and provenance for trust

The trust capability embedded in the architecture is the key distinction between data space interoperability on the one hand, and commercial platform interoperability and traditional industry consortium data exchange on the other. Provider-defined data usage control travels with data across the ecosystem, supported by identity verification and provenance tracking. These mechanisms presuppose underlying digital security guarantees (OECD, 2022^[18]): (i) confidentiality (shared data are not disclosed beyond

authorised recipients); (ii) integrity (data and related rules are not tampered with in transit or at rest); and (iii) availability (the data and the infrastructure are reliably accessible). Without these guarantees, access and usage control mechanisms lack credibility regardless of how they are specified and implemented. Implementations range from machine-readable policies (enforced, for example, by connectors) to other mechanisms such as PETs (e.g. federated learning, multi-party computation and differential privacy).

Identity and access control are necessary but not distinctive. All data-sharing mechanisms provide verifiable identity and authorisation: centralised platforms authenticate users through proprietary identity and access management systems; bilateral contracts rely on legal identification clauses; services based on PETs authenticate participants before granting access to secure enclaves. Data spaces implement the same function through verifiable credentials (IDS-RAM) or OpenID Connect (ODS-RAM). As the key difference, identity is federated or decentralised rather than managed by a single operator. This matters for governance neutrality, since no single party controls the identity layer. However, it does not by itself distinguish data spaces from other mechanisms.

A highly distinctive capability is data usage control: provider-defined, machine-readable or otherwise technically supported conditions on what may happen to data *after* access is granted e.g. purpose limitation, time bounds, geographic restrictions, prohibition of onward sharing. On centralised platforms, the operator's terms govern post-access use; in bilateral contracts, usage conditions lack technical enforcement; in secure computation, control is limited to computation-only use cases. A comparative of data spaces is the ability to provide user-defined, technically enforceable access *and* usage control across a multi-party ecosystem.

This capability addresses a barrier documented across sectors, private and public alike. Regulatory requirements, including data protection regulations, play a critical role in safeguarding individuals and preventing the misuse of data. However, possibly inappropriate interpretations of these requirements can overly constrain data accessibility even where sharing would be lawful and beneficial. Organisations are thus often reluctant to share without assurance that robust safeguards operate in the receiving entity (OECD, 2025^[19]). By making conditions such as purpose limitation explicit, machine-readable and auditable, usage control gives data providers, and the officers responsible for their compliance, demonstrable grounds for approving sharing that over-cautious interpretation would otherwise block.

The primary mechanism proposed for data usage control, in particular by the IDSA-RAM, is through machine-readable conditions expressed in the Open Digital Rights Language (ODRL),¹⁴ negotiated through the Eclipse Dataspace Protocol and enforced by connectors.¹⁵ This faces an inherent limitation: once data leave the provider's infrastructure, enforcement depends on the data user's compliance. In practice, data-sharing agreements complement technical controls. To that end, they provide a contractual fallback enforceable through legal channels when technical enforcement fails or proves infeasible. PETs address this limitation by shifting from controlling access to controlling what can be learnt (Box 3.2).

Architectural variants illustrate different implementation strategies and, in the case of usage control, different design philosophies. The IDS-RAM/Eclipse Dataspace Protocol model mandates the full stack, including a mandatory negotiation layer even for fully permissive contracts (i.e. "no operation" policies that impose no restrictions). The ODS-RAM makes the contracting module optional and does not restrict the technical means for usage control. It provides only interfaces that accommodate diverse methods – from contract negotiation to PETs such as trusted execution environments and differential privacy (Tsuda, 2026^[16]). The rationale is that binding usage control to a single, institutionally derived protocol can create a bottleneck for market adoption. Thus, while the IDS-RAM approach prescribes the mechanism and achieves adaptability through functional configuration, the ODS-RAM prescribes only the interface and delegates the choice of mechanism to participants.

Provenance and audit are key for making data usage control credible despite the post-transfer enforcement gap. The data user's access interface (e.g. connector or other protocol-level endpoints) is both the enforcement mechanism and the potential point of circumvention. Consequently, usage control policies

alone cannot prevent a determined data user from extracting and misusing data outside the enforcement path. Provenance tracking and audit logs serve as the compensating mechanism. While they do not prevent agreement violations, they make violations more detectable, shifting the problem from technical prevention to deterrence through accountability.

Box 3.2. Privacy-enhancing technologies as complement to contract-based usage control

Privacy-enhancing technologies (PETs) enable the collection, processing, analysis and sharing of information, while safeguarding data confidentiality and privacy (OECD, 2023^[20]; OECD, 2025^[19]). The literature identifies four PET categories as particularly relevant to data spaces:

- Federated learning enables model training across participants without data leaving the provider's infrastructure. This has been demonstrated in healthcare settings where institutions across multiple countries collaboratively trained diagnostic models while preserving privacy.
- Secure multi-party computation allows participants to jointly compute functions over combined data without any party revealing individual inputs.
- Trusted execution environments, often considered as the foundation of confidential computing, provide hardware-isolated enclaves where data can be processed in decrypted form but cannot be observed or extracted by external software. Closely related architectures, such as Data Control Rooms, package these enclaves with policy enforcement and audit capabilities into a managed environment in which approved analyses only run on the provider's infrastructure.
- Synthetic data generation produces artificial datasets that preserve the statistical properties of original data while containing no real records. This enables participants to share representative data for model development, testing or benchmarking without exposing sensitive inputs.

Recent literature suggests that combining PETs with the contract-based framework is more effective than deploying either in isolation. In this approach, a data space connector negotiates a contract that specifies not only who may access data and for what purpose, but through which PET the data must be processed. This makes the choice of privacy technology an enforceable element of the data-sharing agreement. The provider no longer needs to trust the recipient's compliance because the technology structurally limits what can be done with the data. PET adoption in data spaces does, however, face the same collective-action challenges identified throughout this section: additional computational costs, specialised expertise requirements and the need for co-ordination among participants on which technologies to deploy (Auñón et al., 2024^[21]).

The adoption data reflect this shift from prevention to deterrence. While 35% of all initiatives have adopted the access and usage policy building block, only 28% have done so at operational stage. Meanwhile, the provenance building block is moving in the opposite direction – from 25% to 48% (see Figure 3.1). The divergence becomes intelligible when usage control is understood as a capability embedded in the interoperability stack rather than a stand-alone adoption driver. Participants deploy connectors for interoperability and usage control comes bundled. Provenance must be deliberately adopted as data spaces mature and the stakes of policy violation increase.¹⁶ The capability to enforce usage conditions, even when not actively exercised, functions as embedded option value. This is a theoretically grounded hypothesis, not an empirically established finding.

3.4. The bootstrapping problem as central barrier to viability and adoption

Data spaces may address the first-order co-ordination problem identified in Section 3.1, but establishing them faces a second-order collective-action challenge of its own, commonly referred to as the

bootstrapping problem.¹⁷ Data space components, including connectors, semantic models and trust frameworks, are costly to build, but – once designed – they can be re-used by each new entrant.¹⁸ The shared infrastructure layer exhibits club good properties: high fixed costs, low marginal costs for each additional deployment and benefits that are non-rivalrous within the ecosystem but excludable to non-participants (Brousseau, Toledano and Eustache, 2024^[8]). Individual participants, however, may face onboarding costs (e.g. adapting internal systems, implementing connectors, mapping data to shared semantic models) that can be substantial, particularly for small and medium-sized enterprises (SMEs) (DSSC, 2025^[22]).

Initiatives such as the Dataspace Adoption Forum (DSAF) have emerged to help organisations navigate onboarding costs.¹⁹ The bootstrapping challenge therefore operates at two levels: the collective investment in shared infrastructure must precede the network effects that justify it; and each participant's private onboarding investment must be justified by expected returns that may not materialise until the ecosystem reaches sufficient scale.²⁰

The free-riding problem identified by Olson (1965^[1]) applies directly, since participants who can benefit from the shared infrastructure without contributing to its establishment have no rational incentive to invest in its provision. Brousseau, Eustache and Toledano (2024^[8]) argue that multi-sided data ecosystems without a dominant key actor face the structurally hardest co-ordination problem, requiring neutral intermediation that commercial platforms cannot provide. Bisière et al. (2025^[23]) formalise the resulting chicken-and-egg dynamic: data spaces need participants to create value, but participants will not join without existing value. Moreover, the neutrality requirements of the EU DGA further constrain the cross-subsidisation strategies that commercial platforms use to bootstrap participation (Rochet and Tirole, 2003^[24]). The bootstrapping problem is accordingly invoked to justify public investment in data space establishment (Bisière et al., 2025^[23]; Otto and Tardieu, 2025^[25]) (see Chapter 4).

Frischmann (2012^[4]), however, argues that free riding does not necessarily reduce the investment incentives of contributing participants. A participant who benefits from shared infrastructure without contributing proportionally is not undermining the investment case if the contributing participants capture enough private value from their own participation to justify their costs regardless.²¹ This implies that bootstrapping is a binding constraint specifically when the expected private returns to anchor participants fall below their investment threshold despite positive collective returns (i.e. when no subset of participants can justify the investment from their own expected returns alone).

The more recent economics literature identifies several strategies for overcoming the chicken-and-egg dynamic (Caillaud and Jullien, 2003^[26]; Evans and Schmalensee, 2010^[27]). Such strategies include attracting participants who can act on both sides of the market (as both data providers and consumers). This would help reduce the need to build each side independently. Instead existing platforms or ecosystems can be leveraged to converge activities into a common infrastructure; high-value initial use cases can demonstrate tangible benefits to early adopters (a “seeding strategy”); “marquee users” can be secured whose participation signals credibility and attracts others; and one side of the market can subsidise the other during the early adoption phase.

The OECD Expert Workshop confirmed that data space practitioners draw on this repertoire (Annex A). A further approach involves keeping the data space operator deliberately compact by connecting “collaborative services” (shared infrastructure guided by public guidelines) with “competitive services” offered by third-party providers through open application programming interfaces (APIs). This approach limits the operator's own cost structure while enabling a broader ecosystem of value-added services (Annex A). The bootstrapping threshold is further inflated by a recurring co-ordination cost that practitioners describe as “reinventing the wheel” (Annex A).

Data spaces are frequently developed by consortia whose members vary from one initiative to another, bringing different champions, expertise and views. Excessive time spent debating technical components from scratch is a recurring pattern, slowing development and increasing costs. Baron, Meniere and

Pohlmann (2014^[28]) analyse a structurally similar phenomenon in technology standardisation, where formal processes entail duplicative research and development, as well as delays due to vested interests. A key benefit of industry consortia is precisely to reduce these costs by settling conflicts of interest upfront. The data space ecosystem, with multiple overlapping initiatives (IDSA, Gaia-X, European Data Spaces Support Centre, sector-specific consortia) each developing partially overlapping building blocks, faces analogous risks.

The Finnish MDS overcame the bootstrapping problem (Box 3.3). Fintraffic, a state-owned enterprise with an institutional mandate, serves as anchor participant. Competitors (shipping lines, port operators) achieved quantifiable cost reductions in fuel, port fees and emissions penalties thanks to their participation in the data space. Operational trials of the Virtual Port Arrivals use case between Lulea and Oxelosund achieved an average 24% reduction in CO₂ emissions and 20% reduction in energy consumption (ELS Shipping, n.d.^[29]). Extrapolated to all Finnish maritime traffic, the projected annual savings amount to EUR 45 million in fuel and costs related to the EU Emissions Trading System (2022 baseline), as well as 135 000 metric tonnes of avoided CO₂ emissions (Siili Solutions, n.d.^[30]).²² A concrete revenue-sharing model distributes these savings among the shipping company, the charterer, the port and the system operator (Annex A). The private return to participants exceeds their participation cost; bootstrapping was not a binding constraint because the anchor institution and the quantifiable value proposition together cleared the investment and participation threshold.

Box 3.3. The case of the Maritime Data Space

Fintraffic is a state-owned enterprise in Finland tasked with the management and development of the nation's traffic control services. Through its Maritime Data Space (MDS) project, Fintraffic develops practical solutions for sharing data generated by ships. The initiative aims to optimise traffic, reduce fuel consumption, lower emissions and address numerous logistical challenges within ports. At the same time, it seeks to improve safety by sharing real-time information about ship movements, weather conditions and other critical safety factors.

Contrary to previous maritime logistics initiatives that focussed on specific company issues, MDS aims to deliver value to the entire ecosystem. Cost sharing among various industry players is a key value proposition centred around two use cases to optimise ship scheduling and fuel efficiency, thereby reducing delays, costs and environmental impacts. In Just-In-Time Arrivals, data analytics optimise a ship's journey and port visit by considering factors such as weather, cargo specifics and timings for arrivals, departures and turnarounds. Overall, it aims to enhance the cost effectiveness of shipping operations and the use of port resources. For its part, Virtual Port Arrivals aims to optimise the arrival process of ships at ports by allowing ships to adjust their arrival speed in real time based on the port's current use rates.

A key factor contributing to the financial viability and success of MDS has been attributed to its focus on governance, business models and value creation prior to technological implementation. The OECD Expert Workshop identified five sequential steps in this approach (Annex A): (i) identify ecosystem champions from shipping companies, charterers, ports, cargo owners and relevant public entities; (ii) develop systemic service design by mapping stakeholders' challenges and incentives, their roles and operations, data holdings and flows, and existing contractual frameworks; (iii) map the real-life contractual model and value model; (iv) build the business model, including a revenue-sharing arrangement among participants; and (v) enable the solution through data space technology. This process took approximately eight months. The emphasis on demonstrating tangible monetary savings alongside environmental benefits proved essential for securing industry buy-in: financial advantages and environmental compliance aligned in the same direction.

Agdatahub, which declared bankruptcy in November 2024 (Agdata, 2024^[31]), represents the opposite case of the bootstrapping problem (Box 3.4). The value of shared agricultural data remained speculative; no anchor participant had a private return sufficient to justify the investment and participation costs on a continuous basis. The data space seemed to have been designed around funding availability rather than an identified business problem. When public grant funding from the French government expired, participation could not sustain itself.

The OECD Expert Workshop highlighted a structural variable: the agricultural sector's fragmentation into many SMEs made it inherently harder to organise into a data-sharing ecosystem than sectors with concentrated global companies sharing common supply chains (Annex A). This is consistent with the economics literature suggesting that large, fragmented groups face stronger free-rider problems than small, concentrated ones; the individual payoff from collective provision decreases with group size, while the cost of organising voluntary contributions increases.²³

In markets dominated by SMEs, regulatory requirements may not serve as an effective incentive for participation. SMEs often need to digitise their operations first, and additional regulatory compliance requirements only increase costs and reduce their capacity to innovate with shared data (Annex A). Similar digitalisation constraints apply beyond firms, such as to government ministries and public administrations (including municipalities) whose operations are often not yet fully digitised.

Box 3.4. The case of Agdatahub

Founded in 2017 by the French agricultural sector with the support of technical institutes and chambers of agriculture, Agdatahub launched Agritrust around 2020. It received significant additional funding from the French government's *France 2030* initiative (Agdata, 2024^[31]). Agdatahub operated as a trusted data intermediary and data space operator for the French and European agricultural sectors, aiming to facilitate secure data sharing between farms and providers of agricultural technologies and services. Its initial goal was "to provide the 380 000 farms in France and all the relevant actors in the sector with innovative tools that take into account the consent of farmers to the use of their data" (Avril, 2023^[32]).

Agdatahub's business model focussed on providing digital tools for managing agricultural data while ensuring compliance with European regulations. This included two flagship offerings (Avril, 2023^[32]): Agritrust and API-Agro. Agritrust was conceived as a solution for managing farmers' consent to the use of their data, based on "the very first agricultural digital identity". Focusing on digital service providers for farms and available as a paid subscription, Agritrust was also provided as a free mobile application for farmers. API-Agro was a platform for multi-stream exchange between holders and users of agricultural data. The subscription revenue model (including a first level free of charge) allowed relevant sectoral actors to access numerous functionalities and exchange data securely.

Both solutions were designed to meet the needs of the agricultural industry in terms of data dissemination for innovations; to address data control and transparency requirements weighing on farmers; and to respect EU regulations, including the EU DGA and Data Act. Despite its key objective to promote innovation for private stakeholders on the agriculture market, Agdatahub stated its business model proved unsustainable without continued government funding.

BoostAeroSpace occupies an intermediate position as a governance consortium that illustrates lessons relevant to data spaces. Major original equipment manufacturers in the aerospace sector captured enough private return from supply chain efficiency gains to anchor the investment, even as smaller suppliers free-rode on the shared infrastructure. The trajectory from founding (approximately 2011) to financial viability (approximately 2019) holds two key lessons: bootstrapping can be overcome through sufficient anchor participant returns; and the time horizon required can be considerably longer than a typical public grant cycle of two to four years. This has implications as discussed in Chapter 4.

A common pattern emerges across the three cases. The actual value in data spaces is created at the level of individual use cases, not at the level of the infrastructure itself (Annex A). The infrastructure layer provides the common governance, interoperability and trust framework within which use cases operate. However, participants invest because of the concrete value they expect to derive from specific applications. In the cases discussed, and confirmed more broadly at the OECD Expert Workshop across practitioners from Europe, Japan and North America, prioritising use-case value over infrastructure development was consistently associated with stronger prospects for financial sustainability (Annex A). The Finnish MDS designed its business model around two specific use cases before selecting technology. ABtC followed a similar pattern in Japan, anchoring the ecosystem around a concrete initial use case (carbon footprint data for battery compliance) before expanding scope. Agdatahub invested in a general-purpose infrastructure platform without demonstrated use-case value sufficient to sustain participation. The distinction between use-case business models and infrastructure business models is therefore analytically consequential for the policy analysis in Chapter 4.

The three cases differ systematically along two structural dimensions that prior work on data-sharing ecosystems identifies as critical for viability: (i) whether participants hold symmetric or asymmetric positions in the value chain; and (ii) whether a “keystone actor”²⁴ co-ordinates the ecosystem (Iansiti and Levien, 2004^[33]; Eustache, Brousseau and Toledano, 2025^[34]).

Table 3.2 situates the three cases discussed above, together with the Ouranos battery passport use case (Chapter 4), within a 2×2 matrix. The configurations align with the bootstrapping outcomes observed: where a keystone actor absorbs initial integration costs and asymmetries are managed through redistribution (BoostAeroSpace; Ouranos battery passport), bootstrapping is more readily overcome. Where no keystone is present and participants are heterogeneous in capabilities and incentives (Agdatahub), the orchestrator must perform both strategic alignment and technical intermediation, and the bootstrapping problem is most acute.

Table 3.2. Governance configurations of data spaces based on type of actors

	Keystone present	Keystone absent
Symmetric (shared interests, comparable positions)	Network efficiency under hierarchical co-ordination; risk of monopolised value capture. Illustrative case: Ouranos battery passport use case (diffusion phase).	Collective problem-solving via neutral intermediary; emergence spontaneous but strategic growth hindered by absence of leadership. Illustrative case: Finnish Maritime Data Space.
Asymmetric (heterogeneous positions and capabilities)	Value redistribution; keystone subsidises weaker participants to align incentives. Illustrative case: BoostAeroSpace.	Hardest configuration; orchestrator must combine strategic alignment with technical intermediation without a power base. Illustrative case: Agdatahub (declared bankruptcy 2024).

Source: Author’s elaboration based on Eustache, Brousseau and Toledano (2025^[34]), *The Economics of Data-Sharing: An Empirical Investigation of Data Sharing Ecosystems*, <https://hal.science/hal-05542222v1>; and case studies from Sections 3.3 and 3.4.

A distinct revenue model is emerging from IT service providers, including telecommunications operators and potentially cloud platforms. In this model, data space connectivity is offered as an incremental managed service on top of existing infrastructure rather than building a stand-alone intermediary that must reach its own critical mass (Box 3.5). The model sidesteps the bootstrapping problem by leveraging infrastructure and customer relationships that already exist at scale. It illustrates that commercial viability in data spaces may be achieved by operators with existing infrastructure, neutral intermediary positioning and volume-based revenue that scales with adoption. Whether integrated platform operators can credibly serve this role depends on two factors: the regulatory neutrality requirements discussed in Section 4.1 of Chapter 4; and whether participants will trust an intermediary that also competes for their compute, analytics or AI workloads with their competitively sensitive data.

Box 3.5. Telecommunications operators as data space service providers: The case of T-Systems

T-Systems, the IT services subsidiary of Deutsche Telekom, treats data space connectivity as a managed service comparable to Internet or voice connectivity. Its Data Intelligence Hub offers commercialised products for ecosystem onboarding, sandbox testing, identity verification (based on Gaia-X Digital Clearing House standards) and geo-information analytics. T-Systems was the first organisation to receive IDSA connector certification (January 2024) and serves as a certified onboarding service provider for Catena-X. Revenue is product-based rather than tied to data transactions. The model builds on existing network infrastructure, making data space connectivity an incremental service rather than a stand-alone investment requiring its own critical mass.

The model has a cross-border dimension. In September 2024, T-Systems and the NTT Group (NTT Communications, renamed NTT DOCOMO Business in July 2025, with NTT DATA) deployed a data space test environment in Japan. This integrated Gaia-X trust anchor mutual recognition between Europe and Japan, building on a 2022 NTT–Catena-X partnership announcement (NTT, 2022^[35]; 2025^[36]). IDSA established a Global Telecom Community in June 2024 with T-Systems, NTT Communications (now NTT DOCOMO Business) and KPN (Netherlands) as founding members. NTT DATA has set a target of approximately EUR 310 million in data space-related revenue by 2030, one of the few publicly stated commercial benchmarks in the ecosystem.

References

- Agdata (2024), “Agdatahub demande l’ouverture d’une procédure collective – Agdatahub”, [31]
<https://agdatahub.eu/en/agdatahub-procedure-collective-liquidation/> (accessed on
 25 February 2025).
- Andresel, M. et al. (2024), “Adapting ontology-based data access for data spaces”, The Second [45]
 International Workshop on Semantics in Dataspaces, co-located with the Extended Semantic
 Web Conference, May 26 – 27, 2024, Hersonissos, Greece: SDS 2024 co-located with
 ESWC 2024, <https://ceur-ws.org/Vol-3705/paper06.pdf>.
- Arbin, K. and U. Essler (2005), “Covisint in Europe: Analysing the B2B auto e-marketplace”, [5]
International Journal of Automotive Technology and Management, Vol. 5/1, p. 31,
<https://doi.org/10.1504/ijatm.2005.006878>.
- Arrow, K. (1962), “Economic Welfare and the Allocation of Resources for Invention”, in *The Rate* [44]
and Direction of Inventive Activity, National Bureau of Economic Research, Princeton
 University, <https://doi.org/10.1515/9781400879762-024>.
- Auñón, J. et al. (2024), “Evaluation and utilisation of privacy enhancing technologies – A data [21]
 spaces perspective”, *Data in Brief*, Vol. 55, p. 110560,
<https://doi.org/10.1016/J.DIB.2024.110560>.
- Avril (2023), “Agdatahub raises €4.8 million to boost the development of a French and European [32]
 agricultural infratech in a trusted framework”, [https://presse.avril.com/agdatahub-raises-e4-8-](https://presse.avril.com/agdatahub-raises-e4-8-million-to-boost-the-development-of-a-french-and-european-agricultural-infratech-in-a-trusted-framework/?lang=en)
[million-to-boost-the-development-of-a-french-and-european-agricultural-infratech-in-a-trusted-](https://presse.avril.com/agdatahub-raises-e4-8-million-to-boost-the-development-of-a-french-and-european-agricultural-infratech-in-a-trusted-framework/?lang=en)
[framework/?lang=en](https://presse.avril.com/agdatahub-raises-e4-8-million-to-boost-the-development-of-a-french-and-european-agricultural-infratech-in-a-trusted-framework/?lang=en) (accessed on 25 February 2025).
- Baron, J., Y. Ménière and T. Pohlmann (2014), “Standards, consortia, and innovation”, [28]
International Journal of Industrial Organization, Vol. 36, pp. 23-35,
<https://doi.org/10.1016/j.ijindorg.2014.05.004>.
- Battaglini, M. and T. Palfrey (2024), “Organizing for collective action: Olson revisited”, *Journal of* [43]
Political Economy, Vol. 132/9, <https://doi.org/10.1086/729580>.
- Bisière, C. et al. (2025), “The economics of data spaces”, *Policy Paper*, SE Digital Center and [23]
 Toulouse School of Economics, France, [https://www.tse-](https://www.tse-fr.eu/sites/default/files/TSE/documents/DigitalCenter/policy_paper/the_economics_of_data_spaces_july_2025_policy-paper.pdf)
[fr.eu/sites/default/files/TSE/documents/DigitalCenter/policy_paper/the_economics_of_data_s](https://www.tse-fr.eu/sites/default/files/TSE/documents/DigitalCenter/policy_paper/the_economics_of_data_spaces_july_2025_policy-paper.pdf)
[paces_july_2025_policy-paper.pdf](https://www.tse-fr.eu/sites/default/files/TSE/documents/DigitalCenter/policy_paper/the_economics_of_data_spaces_july_2025_policy-paper.pdf).
- Brousseau, E., J. Toledano and L. Eustache (2024), “Economics of data sharing: Study on the [8]
 emergence and creation of value within data-sharing ecosystems”, *Position Paper*, Paris
 Dauphine University / Gaia-X Institute, [https://gaia-x.eu/wp-content/uploads/2024/03/Study-](https://gaia-x.eu/wp-content/uploads/2024/03/Study-on-the-emergence-and-creation-of-value-within-data.pdf)
[on-the-emergence-and-creation-of-value-within-data.pdf](https://gaia-x.eu/wp-content/uploads/2024/03/Study-on-the-emergence-and-creation-of-value-within-data.pdf).
- Caillaud, B. and B. Jullien (2003), “Chicken & egg: Competition among intermediation service [26]
 providers”, *The RAND Journal of Economics*, Vol. 34/2, pp. 309-328,
<https://doi.org/10.2307/1593720>.
- Catena-X (2025), “Catena-X: 10 golden rules”, [https://catenax-ev.github.io/docs/next/regulatory-](https://catenax-ev.github.io/docs/next/regulatory-framework/30000ft/ten-golden-rules) [14]
[framework/30000ft/ten-golden-rules](https://catenax-ev.github.io/docs/next/regulatory-framework/30000ft/ten-golden-rules) (accessed on 10 April 2026).
- Catena-X (2023), “About Us”, <https://catena-x.net/en/about-us> (accessed on 10 April 2026). [13]

- DSSC (2026), “Data Spaces Blueprint v3.0”, [10]
<https://dssc.eu/space/BVE/357074549/Data+Spaces+Blueprint+v3.0> (accessed on 9 April 2026).
- DSSC (2025), “Participant adoption and sustainability in data spaces”, [22]
<https://dssc.eu/publications> (accessed on 16 April 2026).
- ELS Shipping (n.d.), *Virtual Arrival*, <https://www.eslshipping.com/en/services/virtual-arrival> [29]
 (accessed on 14 October 2026).
- Eustache, L., E. Brousseau and J. Toledano (2025), “The economics of data-sharing: An empirical investigation of data sharing ecosystems”, *Forty-Sixth International Conference on Information Systems (ICIS)*, Nashville, Tennessee, <https://hal.science/hal-05542222v1>. [34]
- Evans, D. and R. Schmalensee (2010), “Failure to launch: Critical mass in platform businesses”, [27]
Review of Network Economics, Vol. 9/4, <https://doi.org/10.2202/1446-9022.1256>.
- Frischmann, B. (2012), *Infrastructure: The Social Value of Shared Resources*, Oxford University Press. [4]
- Gelhaar, J. et al. (2021), “Towards a taxonomy of incentive mechanisms for data sharing in data ecosystems”, *PACIS 2021 Proceedings*, ACIS 2021 Proceedings 121, [7]
<https://aisel.aisnet.org/pacis2021/121/>.
- Gerst, M. and R. Bunduchi (2005), “Shaping IT standardization in the automotive industry – the role of power in driving portal standardization”, *Electronic Markets*, Vol. 15/4, pp. 335-343, [6]
<https://doi.org/10.1080/10196780500302872>.
- Guggenberger, T., C. Schlueter Langdon and B. Otto (2025), “Data spaces as meta-organisations”, *European Journal of Information Systems*, Vol. 5/34, pp. 822-842, [11]
<https://doi.org/10.1080/0960085X.2025.2451250>.
- Hart, P. and C. Saunders (1997), “Power and trust: Critical factors in the adoption and use of electronic data interchange”, *Organization Science*, Vol. 8/1, [42]
<https://doi.org/10.1287/orsc.8.1.23>.
- Iansiti, M. and R. Levien (2004), “The Keystone advantage: What the new dynamics of business ecosystems mean for strategy, innovation, and sustainability”, *Academy of Management Perspectives*, Vol. 20/2, <https://doi.org/10.5465/AMP.2006.20591015>. [33]
- IDSA (2024), *IDS Reference Architecture Model, Version 4.0*, https://github.com/International-Data-Spaces-Association/IDS-RAM_4_0. [41]
- IPA (2026), “Open Data Spaces Reference Architecture Model (ODS-RAM)”, <https://open-dataspaces.gitbook.io/ods-docs/ods-ram/v2> (accessed on 1 April 2026). [15]
- IPA and Catena-X (2025), “Catena-X and Ouranos ecosystem successfully demonstrate data space interoperability”, <https://catena-x.net/news/catena-x-and-ouranos-ecosystem-successfully-demonstrate-data-space-interoperability/> (accessed on 6 March 2026). [17]
- Kernstock, P. et al. (2025), “Establishing and governing data ecosystems at the crossroads of centralization and decentralization”, *Electronic Markets*, Vol. 35/71, [12]
<https://doi.org/10.1007/s12525-025-00810-x>.

- Löhe, J. and C. Legner (2010), “SOA adoption in business networks: do service-oriented architectures really advance inter-organizational integration?”, *Electronic Markets*, Vol. 20/3-4, pp. 181-196, <https://doi.org/10.1007/s12525-010-0046-7>. [40]
- NTT (2025), “Renewal of NTT Group’s corporate identity”, <https://group.ntt/en/newsrelease/2025/05/09/250509z.html> (accessed on 4 May 2026). [36]
- NTT (2022), “NTT Com and NTT DATA to develop data-sharing eco-system that interconnects with Europe’s Catena-X Platform to protect data sovereignty will enable secure sharing of automotive, aviation & energy data in global value chains”, <https://group.ntt/en/newsrelease/2022/05/26/220526b.html> (accessed on 4 May 2026). [35]
- OECD (2025), “Sharing trustworthy AI models with privacy-enhancing technologies”, *OECD Artificial Intelligence Papers*, No. 38, OECD Publishing, Paris, <https://doi.org/10.1787/a266160b-en>. [19]
- OECD (2023), “Emerging privacy-enhancing technologies: Current regulatory and policy approaches”, *OECD Digital Economy Papers*, No. 351, OECD Publishing, Paris, <https://doi.org/10.1787/bf121be4-en>. [20]
- OECD (2022), “Measuring the value of data and data flows”, *OECD Digital Economy Papers*, No. 345, OECD Publishing, Paris, <https://doi.org/10.1787/923230a6-en>. [46]
- OECD (2022), *OECD Policy Framework on Digital Security: Cybersecurity for Prosperity*, OECD Publishing, Paris, <https://doi.org/10.1787/a69df866-en>. [18]
- OECD (2019), *Public Value in Public Service Transformation: Working with Change*, OECD Publishing, Paris, <https://doi.org/10.1787/47c17892-en>. [9]
- Olson, M. (1965), *The Logic of Collective Action: Public Goods and the Theory of Groups*, Harvard University Press. [1]
- Otto, B. and H. Tardieu (2025), *Data Sharing: Closing the Gap in the Draghi Report’s Call for AI Competitiveness in Europe*, Le Grand Continent. [25]
- Rochet, J. and J. Tirole (2003), “Platform competition in two-sided markets”, *Journal of the European Economic Association*, Vol. 1/4, pp. 990-1029, <https://doi.org/10.1162/154247603322493212>. [24]
- Siili Solutions (n.d.), “Virtual port arrival - maritime innovation”, *Siili Solutions*, <https://www.siili.com/cases/virtual-port-arrival-maritime-innovation> (accessed on 10 April 2026). [30]
- Staebler, M., F. Koester and C. Schlueter-Langdon (2023), “Towards solving ontological dissonance using network graphs”, *Americas Conference on Information Systems*, p. 2023, <https://arxiv.org/abs/2308.14326v1>. [39]
- Swink, M. et al. (2024), “Supply chain visibility types and contextual characteristics: A literature-based synthesis”, *Journal of Business Logistics*, Vol. 45/1, pp. E-12366, <https://doi.org/10.1111/jbl.12366>. [38]

- Tsuda, M. (2026), *Why Open Dataspaces: Design Philosophy and Architectural Paradigm*, Information-technology Promotion Agency, Tokyo, [16]
https://www.ipa.go.jp/en/digital/architecture-guidelines/individual-link/p1o1lf000001xv4n-att/WhyOpenDataspaces_en.pdf.
- Williamson, O. (1991), "Comparative economic organization: The analysis of discrete structural alternatives", *Administrative Science Quarterly*, Vol. 36, pp. 269-296, [3]
<https://doi.org/10.2307/2393356>.
- Williamson, O. (1985), *The Economic Institutions of Capitalism: Firms, Markets, Relational Contracting*, The Free Press, New York. [2]
- World Wide Web Consortium (2018), "ODRL Information Model 2.2", W3C, [37]
<https://www.w3.org/TR/odrl-model/> (accessed on 23 April 2026).

Endnotes

¹ The problem is more fundamental than context-dependency alone. Arrow (1962^[44]) identified the information paradox: a prospective buyer cannot assess the value of information without first possessing it, yet once disclosed the information has effectively been transferred. Price negotiation requires revealing the very asset being sold, undermining the conditions for market exchange. In multi-party settings, this paradox is compounded because the same data have different value to each participant, making even bilateral price discovery intractable at ecosystem scale.

² The value depends on what other data it is combined with (combinatorial value), on uses that cannot be fully specified *ex ante* (option value) and on downstream applications whose beneficiaries extend beyond the immediate transaction (demand-side externalities) (OECD, 2022^[46]).

³ Electronic data interchange (EDI) in the 1980s, supplier portals in the 1990s, application programming interfaces and integration platforms in the 2000s, and centralised control towers in the 2010s all enabled data exchange where one dominant party set the terms. This was not the case in multilateral settings where governance had to be shared among participants with competing interests. EDI adoption was characterised by buyer coercion and power asymmetry: dominant firms used EDI to extend control over supply chain partners, and small firms adopted primarily because of external pressure from powerful trading partners rather than perceived mutual benefit (Hart and Saunders, 1997^[42]). Integration based on Service-Oriented Architecture in the 2000s reproduced hub-and-spoke governance dominated by focal organisations despite the technology's promise of dynamic, decentralised business networks; polycentric arrangements between equal partners remained rare (Löhe and Legner, 2010^[40]). Supply chain control towers in the 2010s achieved visibility only within the focal firm's immediate trading partners; governance structure, not technology, is the binding constraint on multi-tier visibility (Swink et al., 2024^[38]).

⁴ Two features of this setting sharpen the problem. Statutory mandates bind each authority's data to defined purposes, so that inter-governmental re-use can require a new legal basis rather than a commercial agreement. In addition, the payoff structure of public administration can institutionalise risk aversion, since officials may face sanctions for wrongful sharing while capturing little of the value that sharing creates.

⁵ The multi-party framing also covers internal data flows within a single global enterprise: cross-country, cross-subsidiary or cross-business-line transfers typically involve semi-autonomous departments, branches or subsidiaries that face the same trust, control and interoperability problems as independent organisations, and increasingly turn to data space architectures to address them.

⁶ The asset-specificity leg requires qualification. Data themselves are nonrival and technically redeployable across relationships at near-zero marginal cost. The specificity lies not in the data but in the participation investment: adapting internal systems; implementing connectors; mapping data to shared semantic models; and building governance relationships are sunk costs whose value is largely specific to the ecosystem (Section 3.4 of Chapter 3). Context-dependent data value reinforces this specificity, since data that are technically portable may yield substantially less value outside the ecosystem where complementary datasets and institutional infrastructure give it meaning. They are, however, appropriate only for a subset of data-sharing problems.

⁷ Catena-X illustrates the interdependency (Box 3.1 of Chapter 3). Distributed governance through Catena-X Automotive Network e.V. establishes the institutional foundation; open interoperability via EDC and IDSA standards provides the technical infrastructure through which competitors actually exchange data; and usage control (the trust capability embedded in the interoperability stack) makes competitors willing to contribute proprietary supply chain data. Regulatory requirements (provenance documentation and identity management for EU Battery Regulation compliance) amplify the urgency (see Chapter 4).

⁸ Society for Worldwide Interbank Financial Telecommunication (SWIFT) operates the global messaging network for financial transactions.

⁹ Société Internationale de Télécommunications Aéronautiques (SITA) provides IT and communications infrastructure for the air transport industry.

¹⁰ GS1 manages global supply chain standards including barcodes and electronic data interchange.

¹¹ In the Finnish Maritime Data Space, quantifiable operational savings (reduced fuel costs, port fees and emissions penalties) are distributed among participants (shipping company, charterer, port and system operator) according to a collectively negotiated formula. In this way, the data space is funded through the value it generates rather than through external grants (see Box 3.3 of Chapter 3 and Annex A).

¹² Each pillar operates independently and is loosely coupled, so participants can selectively implement components based on their use case and maturity rather than adopting the full stack at once. The architecture emphasises “incremental payoff for incremental investment” and “Minimal Yet Viable” design, drawing on the Classical Dataspaces principle that solutions should “not exist only as monolithic solutions”. (Tsuda, 2026^[16]). See also IPA (2026^[15]).

¹³ Ontology matching based on large language models has become an active research area, with dedicated benchmarks (OAEI-LLM, OAEI-LLM-T) and systems such as MILA reporting F1 scores of 0.83-0.95 on biomedical alignment tasks. Within the data spaces context, early work addresses the “ontological dissonance” across participating domains (Staebler, Koester and Schlueter-Langdon, 2023^[39]) and adapts ontology-based data access for data spaces (Andresel et al., 2024^[45]). Within ODS-RAM v2, a “dynamic ontology” concept has been proposed as a design direction (Tsuda, 2026^[16]). Deployment in operational data spaces remains experimental.

¹⁴ ODRL is a W3C Recommendation (World Wide Web Consortium, 2018^[37]) that provides a standardised vocabulary for expressing permissions, prohibitions and obligations over digital content and services. In the data space context, ODRL policies specify conditions such as permitted purposes, temporal validity, geographic restrictions, and onward-sharing prohibitions in a machine-readable format that connectors can interpret and enforce automatically, without requiring human review of each data-sharing request.

¹⁵ The Dataspace Protocol defines a state machine for contract negotiation and data transfer between a provider and a data user connector. The Eclipse Dataspace Components, the most widely deployed reference implementation, operationalises this protocol: the provider connector publishes a catalogue of available data assets with attached ODRL policies; the data user connector initiates a negotiation; once both sides agree on terms, the protocol transitions to a transfer state in which data flow only under the negotiated conditions. Enforcement occurs at the connector level: the provider's connector verifies that the data user has accepted the negotiated conditions before authorising data transfer. Post-transfer enforcement, however, depends on the data user's connector honouring the agreed policies. The IDS-RAM acknowledges that this requires trusted platforms and cannot prevent circumvention by administrators with infrastructure access (Section 4.1.6 of IDS-RAM 4.0) (IDSA, 2024^[41]).

¹⁶ The combination of ODRL policies, connector enforcement and provenance tracking therefore functions as a socio-technical enforcement system: technical controls at the point of transfer, complemented by accountability mechanisms that make post-transfer non-compliance detectable and legally actionable.

¹⁷ The bootstrapping problem in multi-sided ecosystems is analysed in the platform economics literature as a chicken-and-egg dynamic (Caillaud and Jullien, 2003^[26]).

¹⁸ The Eclipse Dataspace Components reference implementation is developed once as open-source software, although each participant must deploy and configure their own instance (managed-service offerings increasingly reduce this burden for smaller firms). An industry-specific semantic model (such as Catena-X's automotive data models) is negotiated once and re-used across the supply chain. A trust framework's certification process, once designed, applies uniformly to new entrants.

¹⁹ These costs reflect the technical complexity of data space infrastructure itself, which some practitioners have described as a barrier to adoption in its own right. Where a need for data sharing involves few parties, bilateral APIs can offer a cheaper and more pragmatic alternative.

²⁰ The concept of "dataspaces" introduced by Franklin et al. eliminates this problem by design, treating integration as a private investment decision in which each participant pays only for connections whose value exceeds their cost. Modern data spaces invert that logic, requiring collective infrastructure to precede individual participation.

²¹ As Frischmann argues, "there is a mistaken tendency to believe that any gain or loss in profits corresponds to an equal or proportional gain or loss in investment incentives, but this belief greatly oversimplifies the decision making process and underlying economics" (Frischmann, 2012^[4]).

²² These figures are based on business calculations from real ESL Shipping voyage data and port call data. The EUR 45 million figure is a counterfactual projection (what would have been saved had the system been operational across Finnish maritime traffic in 2022), not measured aggregate savings. The 24% and 20% reduction figures are from ESL Shipping's operational data on a specific route, not independently verified. No peer-reviewed publication was identified at time of writing.

²³ Battaglini and Palfrey (2024^[43]) formalise this insight using mechanism design. They show that for large groups the probability of successful collective action converges to zero, although optimal organisational mechanisms can slow the rate of decline. No empirical study was identified that applies Olson's group-size logic specifically to data space participation across sectors; the workshop observation represents practitioner experience rather than a tested empirical finding.

²⁴ See lansiti and Levien (2004^[33]) who argue that business ecosystems contain different kinds of actors, which they classify into: (i) Keystones: actors that occupy hubs in the network, set standards and actively shape ecosystem health by increasing value for many others (e.g. Microsoft for the PC ecosystem in the 2000s, Walmart for retail supply chain); (ii) Dominators: actors that exploit their hub position to extract value rather than share it, and often kill the ecosystem they sit in; and (iii) Niche players: specialised actors that depend on the keystone but contribute distinct value.

4

The evolving policy landscape

This chapter reviews the policy instruments through which governments across OECD Member countries and partner economies support data spaces (including functionally comparable architectures). It examines regulatory frameworks that require or encourage multi-party data sharing, the design of public investment on both the supply and the demand side of the bootstrapping problem, standardisation and cross-border interoperability, and certification and public trust frameworks. It concludes with the role of government as data space operator and anchor participant, including inter-governmental data spaces.

Governments across OECD Member countries and Key Partner economies are deploying data spaces and functionally comparable architectures through a widening range of policy instruments. Drawing on examples from all regions, this chapter organises selected policy initiatives by the type of policy intervention, which are often combined. Section 4.1 examines regulatory frameworks that require or encourage multi-party data sharing. Section 4.2 analyses public investment and institutional design. Section 4.3 explores standards and cross-border interoperability, while Section 4.4 looks at certification and public trust frameworks. Section 4.5 examines inter-governmental (G2G) data spaces and government deployment as operator and anchor participant.

4.1. Regulatory frameworks that require multi-party data sharing

Regulatory instruments across jurisdictions are driving data space adoption through two distinct channels:

- Some explicitly mandate or promote data space architectures, as with the European Health Data Space (EHDS) regulation (European Union, 2025^[1]) or India's creation of regulated consent managers under Data Empowerment and Protection Architecture (DEPA) (NITI Aayog, 2020^[2]).
- Others impose requirements that can most readily be met through federated, governed data exchange without specifically targeting data spaces, as with the EU Battery Regulation (European Union, 2023^[3]), or the Trusted Exchange Framework and Common Agreement (TEFCA) in the United States (HHS, 2026^[4]).

In the European Union (EU), several instruments create requirements that align structurally with data space building blocks. The EU Data Governance Act (DGA) (European Union, 2022^[5]) requires separation between data intermediation and other commercial activities. The EU AI Act (European Union, 2024^[6]) imposes provenance and traceability obligations for high-risk AI systems. The EU Data Act (European Union, 2023^[7]) establishes fair access to data from connected products and interoperability requirements. The EU Battery Regulation requires carbon footprint declarations and digital battery passports (DPPs) involving multi-stakeholder data exchange across supply chains. Digital product passports more broadly are an emerging regulatory instrument requiring machine-readable lifecycle data to accompany products; their traceability, provenance and multi-stakeholder access requirements map onto data space infrastructure.

The EU DGA's neutrality requirement (Article 12) deserves particular attention, not least because it is now subject to significant revision. As adopted in 2022, the EU DGA requires data intermediation services to maintain structural (legal) separation between intermediation and other commercial activities. This is a requirement incompatible with the integrated service models of major cloud platforms, which derive revenue from bundling data intermediation with compute, analytics and artificial intelligence (AI) services. For organisations subject to these requirements, data spaces offer a structurally aligned pathway that proprietary platforms do not.

The EU Digital Omnibus proposal (European Commission, 2025^[8]), published in November 2025 alongside the EU Data Union Strategy (European Commission, 2025^[9]), would substantially alter this landscape. The proposal repeals the EU DGA as a stand-alone regulation and consolidates its provisions, together with those of the EU Open Data Directive (European Union, 2019^[10]) and the EU Free Flow of Non-Personal Data Regulation (European Union, 2018^[11]), into the EU Data Act as a single instrument. In the process, two key changes affect data intermediation: registration shifts from mandatory to voluntary (with only registered providers permitted to use a recognition label); and the separation requirement is relaxed from structural (legal) to functional separation. The neutrality protections are retained in principle. Providers, for example, may not exploit data beyond what is needed to operate the intermediation service. However, the shift from structural to functional separation lowers the barrier for integrated platform operators.

Whether these changes strengthen or weaken the regulatory case for data spaces depends on perspective. On the one hand, a lighter data intermediation regime may encourage more providers to enter the market. On the other, it may also erode the distinctive regulatory advantage of data spaces over commercial platforms under the EU DGA's stricter separation requirement. Bisière et al. (2025^[12]) argue that neutrality requirements constrain the cross-subsidisation strategies available to data space operators (the “see-saw effect”). However, they also point out these same requirements prevent commercial platforms from serving the data intermediation function.

The regulatory effect that encourages adoption of data spaces extends beyond the EU. Third-country suppliers and manufacturers must comply with EU product regulations when exporting to the European market. The EU Battery Regulation and the EU Ecodesign for Sustainable Products Regulation (European Union, 2024^[13]) may require these suppliers and manufacturers to adopt or join data spaces. This extraterritorial reach transforms regulatory alignment from a European policy concern into a global market access requirement.

While discussions on data spaces tend to focus on EU developments, they have global relevance. The most striking evidence comes from Japan, where the Automotive and Battery Traceability Centre (ABtC) was driven explicitly by the EU Battery Regulation. Thus, Japanese manufacturers adopted data space principles (federation, provider control, common data models) not because of any domestic policy mandate but rather because compliance with EU regulations required precisely the multi-stakeholder traceability infrastructure that data spaces provide (Box 4.1). The initiative is built on the Open Data Spaces Reference Architecture Model (ODS-RAM) and is less than two years old; its long-term viability remains untested. Project Hyperion, a Dutch initiative at the implementation stage, applies the data space model to DPPs in aviation maintenance. It shares aircraft component data across civil and military supply chains through a trusted data space (iSHARE/FIWARE stack). The EU Ecodesign for Sustainable Products Regulation extends DPP requirements into broader product categories.

In the health sector, regulation is the primary catalyst rather than a secondary consideration. The Flemish Health Data Space (implementation stage) is explicitly driven by the EHDS regulation. Besides the EHDS, it lists the EU Data Act, the EU DGA and the EU AI Act as its motivating legal frameworks. The data space exists because the regulatory environment demands a compliant infrastructure for cross-institutional health data sharing. Indeed, the OECD has identified that such an infrastructure requires dedicated governance frameworks that promote data availability, while ensuring privacy protection as called for in the OECD Recommendation on Health Data Governance (OECD, 2017^[14]).

Non-European jurisdictions have arrived at functionally comparable regulatory frameworks through independent institutional pathways. In the United States, TEFCA operates as a federated “network of networks” for health data exchange. Developed by the Assistant Secretary for Technology Policy within the Department of Health and Human Services, and co-ordinated by the Sequoia Project, the network features 11 designated qualified health information networks. Together, they interconnect 14 214 organisations representing over 79 000 unique connections. This entails the exchange of nearly 500 million health records as of February 2026, up from approximately 10 million in January 2025 (HHS, 2026^[4]). Six defined exchange purposes (treatment, payment, healthcare operations, public health, government benefits determination and individual access services) provide purpose-bound, contractually enforceable constraints on data use. TEFCA addresses the same multi-party co-ordination problem that European health data spaces target. However, it does so through a regulatory and contractual framework rather than through a data space architectural stack.

Box 4.1. The case of the Automotive and Battery Traceability Centre (ABtC)

The Automotive and Battery Traceability Centre (ABtC) was established in February 2024 by 14 Japanese automotive manufacturers and two industry associations (the Battery Association for Supply Chain and the Japanese Auto Parts Industries Association). To address the growing global demand for disclosure of information regarding CO₂ emissions and due diligence across the entire supply chain, ABtC operates a traceability service. This enables members to compile total CO₂ emissions (CFP) throughout the battery lifecycle. To that end, they share and link data securely among members within the automotive and battery supply chains, while safeguarding their trade secrets and other sensitive information.

The Centre has four features: (i) ensuring data interoperability between the different applications adopted by its users; (ii) storing data in a federated manner and protecting it using blockchain technology; (iii) data transfer mechanisms using smart contract based on prior agreements; and (iv) the development of common rules for data sharing with ABtC acting as the main co-ordinator (e.g. for defining data models and identifiers).

The value proposition of ABtC is based on enhancing the ability to share data securely and efficiently (at lower costs) across value chains; providing cost-sharing arrangements to reduce compliance burdens, including for small and medium-sized enterprises; and optimising global supply chain management. The revenue model of ABtC is based on a combination of member (basic) fees and transaction (data use) fees (Annex A). It has been accepted to some funding programmes run by the Japanese government for its proof-of-concept experiments, which could eventually be a new service offered in its data space.

ABtC has also obtained accreditation as a “public interest digital platform operator” under a certification framework administered by the Information-technology Promotion Agency (Section 4.4). A notable feature is the sequenced expansion strategy. It begins with carbon footprint traceability for batteries (a “collaborative service” addressing a shared regulatory challenge). Subsequently, it expands to other vehicle components and full vehicles. Eventually, it moves towards “competitive services” such as battery passports and supply chain resilience applications (Annex A). ABtC’s operator has adopted the Open Data Spaces Reference Architectural Model and its protocols for technical standards and data-sharing rules. In so doing, it avoids the need to build its own infrastructure from scratch.

4.2. Public investment and institutional design

Section 3.4 of Chapter 3 established that data spaces face a bootstrapping problem operating at two levels: the collective investment in shared infrastructure, and the individual participant’s onboarding investment. Public policy has responded primarily through two channels: (i) time-limited project funding directed at infrastructure and consortium formation; and (ii) permanent institutional anchoring, in which governments establish or designate organisations with a long-term mandate to co-ordinate data space ecosystems beyond any single funding cycle.

The distinction between the two channels matters. The bootstrapping problem requires sustained co-ordination over time horizons that typically exceed project grant cycles of two to four years (Section 3.4 of Chapter 3). The time-limited funding channel itself operates on two sides of the bootstrapping problem: supply-side investment in shared infrastructure (Section 4.2.1) and demand-side support for participant adoption (Section 4.2.2). The balance between the two has significant implications for whether funded data spaces reach operational viability and scale.

Several jurisdictions have dedicated permanent institutions to provide the long-term co-ordination that project funding cannot. Finland's approach, led by Sitra (the Finnish Innovation Fund), encompasses over 30 data space projects across 15 sectors involving more than 100 organisations (Sitra, 2024^[15]). Sitra has launched the Data Spaces Alliance Finland as a national co-ordination network. It also co-ordinates Towards the European Health Data Space, an EU joint action on the EHDS. In the Netherlands, the Centre of Excellence for Data Sharing and Cloud, co-ordinated by the Netherlands Organisation for Applied Scientific Research, has identified 47 active data-sharing initiatives supported by over 80 service providers; nearly 80% of these initiatives depend on government funding (CoE-DSC, 2024^[16]). In Germany, the cross-ministerial National Data Institute is designed as a hub for co-ordinating data ecosystems across public, private and academic sectors. These institutional arrangements share a common design logic: they separate the co-ordination function from the project-funding cycle.

At the EU level, for instance, the European Data Spaces Support Centre (DSSC) is a consortium of 12 organisations. It is co-ordinated by the Fraunhofer Institute and Research Units and funded under the EU Digital Europe Programme (October 2022 to March 2026). The DSSC provides cross-sectoral co-ordination through the Blueprint, the Community of Practice, and guidance on participant adoption and sustainability (DSSC, 2024^[17]; 2025^[18]). However, the DSSC itself a time-limited project (now continued through a follow-on DSSC 2 phase). This illustrates the tension between the co-ordination function it serves and the project-funding model through which it is financed.¹

4.2.1. Supply-side: Funding data space infrastructure

The dominant policy response to date has been supply-side funding directed at data space infrastructure and consortium formation. As of November 2025, the EU had invested approximately EUR 336 million in Common European Data Spaces across 14 sectors and domains through the Digital Europe Programme. Ongoing EU investment of around EUR 100 million supports their further roll-out (European Commission, 2025^[9]). The EU Data Union Strategy, published in November 2025 alongside the EU Digital Omnibus simplification package (Section 4.1), marks a shift towards implementation and deployment (European Commission, 2025^[9]). The strategy identifies three priority areas: scaling up access to data for AI; streamlining data rules for legal certainty and reduced compliance costs; and safeguarding the EU's position on international data flows. In particular, the strategy introduces "data labs" to link data spaces to AI ecosystems. In this way, it repositions data spaces from a governance instrument to AI supply chain infrastructure.

At the national level, approaches vary significantly in scale and design. Germany committed EUR 200 million to the Gaia-X Funding Competition (Förderwettbewerb Gaia-X), administered by the Bundesnetzagentur. The competition led to 11 selected projects and approximately EUR 117.4 million awarded by end 2024, spanning health, mobility, automotive, energy and manufacturing (Bundesnetzagentur, 2022^[19]). The funding competition phase concluded in March 2025, with the Gaia-X Hub Germany shifting to an onboarding and community co-ordination role. The German approach channels substantial public resources through competitive selection, reflecting the pre-competitive research and development (R&D) transfer model that also produced the IDSA Reference Architecture (Chapter 2).

France has integrated data space funding into its broader France 2030 cloud strategy. To that end, Bpifrance launched a dedicated call for projects with significant impact on structuring French industrial sectors. However, the specific data spaces allocation within France 2030's broader EUR 2.5 billion AI and digital financing envelope has not been publicly broken out (Bpifrance, 2023^[20]).

Canada has launched the first dedicated national data spaces programme outside Europe and Asia. The National Data Spaces Pilot Funding Program, managed by the Digital Research Alliance of Canada, runs from January 2026 to March 2028 with a budget of up to CAD 3.15 million (Digital Research Alliance of Canada, 2025^[21]).

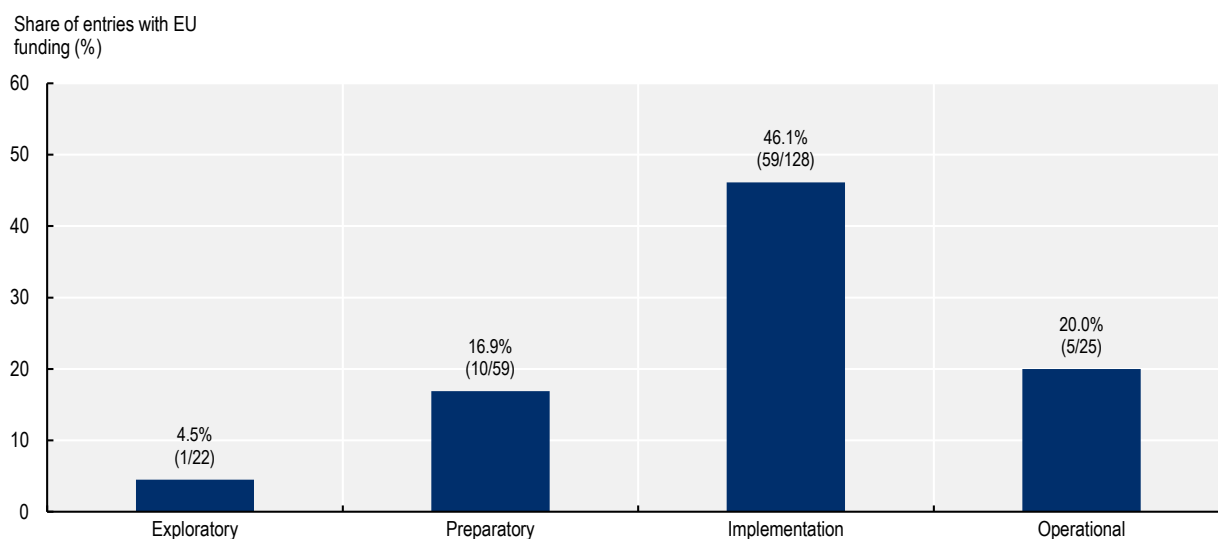
The People's Republic of China (hereafter “China”) has adopted the most explicitly state-directed approach to data spaces. The National Data Administration published a Trusted Data Space Development Action Plan (2024-2028) in November 2024 (National Data Administration, 2024^[22]). It designated 63 pilot projects across enterprise, industry and city levels in 2025, and targets the creation of more than 100 trusted data spaces by 2028.

These supply-side instruments share a structural limitation that becomes clearer when two distinct purposes of public funding are distinguished, as participants at the OECD Expert Workshop emphasised (Annex A). The first purpose is the development of schemas, protocols and interoperability frameworks (i.e. the technical foundation layer). Both public and private entities can host these technical components, and public investment in this layer resembles pre-competitive R&D with broad spillover benefits. The second purpose is the creation of actual data spaces or ecosystems. Here, the case for public funding is more conditional. It depends on whether a specific collective-action problem has been identified that no market actor has the incentive or legitimacy to solve.

EU funding programmes were not explicitly designed around this distinction or the bootstrapping rationale identified in Section 3.4 of Chapter 3. The European Commission's stated framing centres on competitiveness, digital sovereignty and the single market for data, not on correcting a formally identified collective-action failure. The bootstrapping justification is an analytical overlay by economists (Bisière et al., 2025^[12]; Otto and Tardieu, 2025^[23]), not the stated rationale of the policymakers who designed the funding instruments. If the instruments were designed for the first purpose (pre-competitive R&D and technology transfer) rather than the second (solving the specific co-ordination problem faced by data spaces), the fact that only one in nine entries has reached operational status becomes easier to understand. The instruments were simply not designed for the problem they are now expected, or may need, to solve.

The IDSA Radar data are consistent with this diagnosis. EU funding accounts for 46% of implementation-stage entries. However, it drops to 20% at operational stage, where bilateral or multilateral contracts replace grant agreements as the dominant governance model (Figure 4.1) (Section 3.3 of Chapter 3).

Figure 4.1. EU public funding rate by development stage



Note: The public funding rate is the share of entries at each development stage that self-report the EU or its members as a funding source. Funding source is reported by 99 of 235 entries (42%); the remaining entries either did not report or do not have funding. Entries that do not report a funding source are included in the denominator (all entries at that stage), which means the rates shown are lower bounds. EU funding sources include Horizon 2020, Horizon Europe, the Digital Europe Programme, EU4Health and Connecting Europe Facility.

Source: Based on IDSA (n.d.), *IDSA Data Spaces Radar*, www.dataspaces-radar.org.

Yet only 15% of European data-sharing initiatives, including data spaces, have a clear revenue model (Moonen et al., 2025^[24]). The gap between funding prevalence and revenue readiness points to an incentive misalignment in subsidy design (discussed as “moral hazard” in the economics literature).² Where funding instruments reward observable compliance activities (consortium formation, architectural alignment, dissemination) rather than harder-to-measure outcomes (operational viability, participant willingness to pay), rational actors optimise for the former (Holmström and Milgrom, 1991^[25]). Lerner (2009^[26]) documents this dynamic in public venture capital and entrepreneurship programmes. The mechanism – that subsidy design shapes what recipients optimise for – applies by analogy to consortium funding.

The collective-action problem that public funding aims to solve is thus compounded by a second-order problem in which the funding regime itself may generate misaligned incentives. Workshop participants cautioned that long-term public funding may also create dependency effects that prevent data space stakeholders from being efficient and innovative (Annex A). This concern is distinct from the incentive distortion at the margin discussed above; it points to the possibility that the availability of continued public support removes the pressure to find viable revenue models altogether.

The diagnosis is increasingly recognised. Recent work by both practitioner and research communities has placed business value, use cases and sustainable revenue models at the centre of the data spaces agenda (Mertens et al., 2024^[27]; DSSC, 2025^[18]; Moonen et al., 2025^[24]). The DSSC, for instance, notes the focus is shifting “from technical proof of concepts or governance definitions to the practical realities of delivering real value through use cases, attracting, onboarding and retaining participants and reaching a working business model with long-term viability” (DSSC, 2025^[18]). The OECD Expert Workshop held in March 2025 confirmed this finding across practitioners, policymakers and industry experts from Europe, Japan and North America (Annex A).

Prioritising use-case value over infrastructure development was consistently associated with stronger prospects for financial viability. Workshop participants further noted that not all data spaces can be considered public goods. A data space may solve a collective-action problem that no single market actor has the incentive or legitimacy to address (e.g. supply chain traceability across competitors). This solution may create a case for treating the shared infrastructure layer as a public good. However, the dependence of many data spaces on public funding does not in itself demonstrate that they are public infrastructure. It may instead indicate that viable business models have not yet been found.

4.2.2. Demand-side: Funding participant adoption

Far less common is public funding that directly addresses the second level of the bootstrapping problem identified in Section 3.4 of Chapter 3: the onboarding costs that individual participants face when joining a data space. The dominant instruments fund infrastructure (connectors, semantic models, governance frameworks) but not the integration of individual firms into that infrastructure. This gap is significant because even where shared infrastructure exists, participation costs (adapting internal systems, implementing connectors, mapping data to shared semantic models) can deter adoption, particularly by small and medium-sized enterprises (SMEs) (DSSC, 2025^[18]).

Spain’s Data Spaces Kit (*Kit de Espacios de Datos*) is one of the principal exceptions (Red.es, 2025^[28]). Administered by Red.es, Spain’s public entity for digitalisation, the programme subsidises the integration of public and private entities into sectoral data spaces on a non-competitive basis. Drawing on EUR 60 million in NextGenerationEU funds, the Data Spaces Kit provides grants of up to EUR 30 000 for private entities, specifically data providers, and EUR 50 000 for public administrations. The non-competitive, low-threshold design targets breadth of adoption rather than depth of capability. This strategy is suited to building participant density in sectors where the co-ordination problem requires many parties to join simultaneously. The design directly addresses the level-2 bootstrapping problem by reducing the private onboarding cost for each participant, lowering the expected-return threshold that each must clear.

Japan's approach through the Information-technology Promotion Agency (IPA) illustrates a different demand-side instrument: public investment in knowledge infrastructure and capacity building. IPA has developed a structured educational programme (the "Data Spaces Academy") comprising data space guides; a Chief Data Officer handbook for executives on data-driven management; and a data maturity framework targeting C-suite executives, business departments and IT personnel (IPA, 2024^[29]; IPA, 2025^[30]). The rationale, as articulated at the OECD Expert Workshop, is that businesses first need to understand what data spaces are before they can participate; low awareness, limited skills and low data maturity are significant barriers to adoption, particularly among SMEs (Annex A). This capacity-building approach addresses the demand-side bootstrapping problem through a different channel than direct subsidy. Rather than reducing participant onboarding costs financially (as Spain's programme does), it reduces the informational and organisational barriers that prevent potential participants from recognising the value of joining in the first place.

India's DEPA offers a structural alternative to demand-side subsidy (Section 4.1). Rather than funding participant onboarding, the regulatory design creates a category of regulated intermediaries (consent managers) that can charge fees for facilitating data exchange (NITI Aayog, 2020^[2]). This could make the intermediary role itself commercially viable. By presenting a standardised interface to both data providers and data users, consent managers may reduce participant onboarding costs by design rather than by subsidy. India's AA framework has achieved substantial adoption without equivalent public subsidy to the European model. That said, it benefitted from regulatory backing by the Reserve Bank of India and co-ordination by Sahamati, a non-profit industry alliance. The underlying design insight is that creating a commercially viable intermediary category may address the bootstrapping problem more durably than subsidising infrastructure. As such, the design has broader applicability even where the specific consent-based architecture does not transfer directly.

The near-absence of systematic demand-side funding across jurisdictions represents a significant gap in policy design. The three approaches identified above (direct subsidy for onboarding costs, public investment in capacity building and regulatory design that creates commercially viable intermediary categories) each address a different dimension of the level-2 bootstrapping problem. The two-level bootstrapping analysis in Section 3.4 of Chapter 3 suggests that addressing only the supply side (shared infrastructure) while leaving the demand side (participant adoption) unfunded may help explain why data spaces that receive substantial infrastructure investment nonetheless fail to attract sufficient participation.

4.3. Standards, and cross-border interoperability and convergence

Standardisation provides the technical substrate for both domestic and cross-border interoperability. As detailed in Box 2.1 of Chapter 2 (Section 2.2 of Chapter 2), activity is proceeding across multiple bodies: the Eclipse Dataspace Protocol (DSP) is in ISO standardisation as ISO/IEC DIS 26450; CEN/CENELEC and ETSI are developing standards for data space components; and the Institute of Electrical and Electronics Engineers has engaged in related standardisation efforts.

The three major reference architectures (Section 2.2 of Chapter 2) each reflect a different model of public-private collaboration. The IDSA Reference Architecture Model (IDSA-RAM) originated from publicly funded research (approximately EUR 10 million from the German Federal Ministry of Education and Research) carried out by the Fraunhofer Institutes. The resulting model was then transferred to IDSA. ODS-RAM was developed by IPA, a government agency under the Ministry of Economy, Trade and Industry (METI), and published jointly by METI and IPA in February 2025; the Data Society Alliance, a private sector association created to implement government data strategy, co-ordinates ecosystem adoption. The DSSC Blueprint is produced by a consortium of 12 organisations co-ordinated by the Fraunhofer Institutes and Research Units. It is funded by the European Commission under the EU Digital Europe Programme, with the Commission governing at arm's length through the grant framework (see Section 4.2).

Cross-border interoperability between data space ecosystems has advanced furthest between Japan and Europe. Developed under METI with IPA's Digital Architecture Design Center, version 1 of the ODS-RAM was published in February 2025 (METI/IPA, 2025^[31]). In October 2025, IPA, the Data Society Alliance, the Robot Revolution & Industrial IoT Initiative and the University of Tokyo designated "Open Data Spaces" as the common concept for Japan's approach. In April 2026, IPA published the ODS-RAM Version 2 (IPA, 2026^[32]) alongside a companion design philosophy document (Tsuda, 2026^[33]) that elaborates the architectural paradigm (Section 3.3 of Chapter 3). Interoperability with European architecture has been demonstrated in practice: IPA and Catena-X completed a cross-border proof of concept in March 2025 for product carbon footprint data exchange for batteries, driven by EU Battery Regulation compliance requirements (Section 4.1) (METI, 2025^[34]). METI established a certification programme for the Ouranos Ecosystem in May 2025 (METI, 2025^[34]). (See Section 4.4).

Other jurisdictions are building architectures that address the same co-ordination problem, in some cases through independent technical choices. A Green Paper on structuring data spaces in Brazil, prepared by NEO-UFRGS under the Brazilian-German Digital Dialogue, published in July 2024 (NEO-UFRGS, 2024^[35]), outlines eight strategies for developing data spaces in Brazil. It has since established direct links to the European standardisation ecosystem: an IDSA Hub at the Eldorado Research Institute (March 2024, the first in the Americas) and IDSA Competence Center recognition (December 2025). These steps do not yet constitute demonstrated interoperability. However, they align Brazil's emerging data space architecture with the IDSA standards track, reducing the technical distance that future cross-border data exchange would need to bridge.

In the United States, TEFCA achieves federated health data exchange through a common contractual framework rather than through a data space architectural stack (Section 4.1). China's Trusted Data Space Development Action Plan (2024-2028) emphasises blockchain-based traceability, digital contracts with automatic compliance execution and real-time monitoring (National Data Administration, 2024^[22]).

These parallel developments show architectural convergence on common functions (federation, traceability, usage control) without institutional co-ordination. This is a precondition for future cross-border interoperability but not yet evidence of it.

Where cross-border interoperability has been demonstrated, regulatory requirements have created the concrete demand. The Japan-Catena-X proof of concept for battery traceability, driven by EU Battery Regulation compliance, is the clearest case. Standardisation activity provides the technical substrate. So far, however, organisations mainly invest in cross-border interoperability for regulatory reasons rather than market-related rationales (Section 4.1).

Workshop participants drew an analogy with the Internet to argue that standardisation discussions should focus on the protocol layer rather than on how protocols are implemented or used in business (Annex A). The implication is a "coopetition" model in which co-operation on protocols coexists with competition on the services built on top. Identifying what to standardise and keeping it minimal are important design choices. Applied to data spaces, this framing reinforces the analytical distinction between the technical foundation layer and the ecosystem layer identified in Section 4.2.1. It suggests that supply-side public investment may be most productively concentrated at the protocol level. Conversely, demand-side support (Section 4.2.2) remains important for promoting adoption of the services and use cases built on top.

Standardisation approaches across jurisdictions range from the EU's multi-body process through Japan's demonstrated cross-border interoperability to China's state-directed architecture. This diversity constitutes a natural experiment in technical governance design. Its full value depends on systematic measurement, a requirement that current data collection infrastructure does not meet (see Box 1.1 of Chapter 1). The IDSA Radar captures primarily European initiatives; DEPA, TEFCA, ODS and China's trusted data spaces are absent or underrepresented. Extending comparable data collection to non-European federated data-sharing initiatives is a precondition for evidence-based policy guidance that is genuinely international in scope.

4.4. Certification and public trust frameworks

Alongside regulation, public investment and standardisation, certification is emerging as a distinct policy instrument for establishing trust in data space operators and distinguishing those that serve a public interest from those that are purely commercial. Two contrasting approaches are in operation: a registration-based regime and certification.

The EU DGA establishes a registration-based regime for data intermediation services. As adopted in 2022, the regime requires mandatory registration with national competent authorities and structural (legal) separation between intermediation and other commercial activities. The EU Digital Omnibus proposal (Section 4.1) would shift registration from mandatory to voluntary and relax the separation requirement from structural to functional. Under both variants, the mechanism is procedural: operators declare compliance with specified requirements; the competent authority registers (or, under the Omnibus, recognises) them. The regime does not assess the substantive quality or public benefit of the intermediation service, nor does it distinguish between operators serving different purposes or sectors.

Conversely, Japan's "public interest digital platform operator" certification, administered by IPA, certifies operators against substantive criteria for reliability and public benefit. This includes the protection of trade secrets, security, interoperability and data provider control (METI, 2025^[34]). The accreditation of ABtC under this framework (Box 4.1) illustrates how certification can provide institutional legitimacy beyond what a private sector consortium structure alone confers. At the same time, certification also signals to potential participants that the operator has been independently assessed against defined standards. The OECD Expert Workshop discussed this certification approach as a potential model for other jurisdictions seeking to distinguish between data spaces that serve a public interest and those that are purely commercial (Annex A).

The two approaches reflect different policy logics. Registration lowers entry barriers and avoids the administrative cost of substantive assessment but provides limited information to potential participants about operator quality or public benefit. Certification imposes higher entry costs and requires institutional capacity to assess operators but generates a trust signal. This signal may help address the bootstrapping problem (Section 3.4 of Chapter 3) by reducing participants' uncertainty about operator reliability. Whether certification-based or registration-based frameworks prove more effective at establishing durable trust in data intermediaries remains an open empirical question. However, the two approaches offer complementary design insights for jurisdictions developing their own policy frameworks.

Government-backed trust frameworks for digital identity can complement these instruments. For example, the European Digital Identity Wallet – established under the EU's revised eIDAS Regulation (European Union, 2024^[36]) – provides the identity layer on which data space participation and usage control depend (Section 3.3.3 of Chapter 3). This underlines the role that governments can play as an anchor participant in data spaces discussed in Section 4.5.

4.5. Inter-governmental (G2G) data spaces and government as operator and anchor participant

A further policy channel, distinct from regulation, funding, standardisation and certification, is government deployment of data spaces for public sector operations. These configurations span a spectrum. At one end are ecosystems bounded entirely within the public sector to optimise state operations (inter-governmental, G2G). These include cross-ministry planning, public health monitoring or integrated social services. At the other end are arrangements that extend outward to businesses and citizens (G2G2B). In the middle are mixed ecosystems in which government anchors participation among private actors (see Box 4.2). The examples in this section cover this full range.

Rather than subsidising private actors to overcome collective-action barriers, government as operator creates the infrastructure directly and generates the initial data flows and network effects that make subsequent participation attractive. This allows governments to lead by example. By operating data spaces for their own functions, governments demonstrate that the architecture works at scale, legitimise the underlying standards and lower the perceived adoption risk for private actors considering participation. In so doing, government can also address the bootstrapping problem (Section 3.4 of Chapter 3) through a different mechanism than public funding. In some cases, this can lower onboarding costs for subsequent participants.

The choice between funding private data spaces and operating public ones is not binary. The most effective configurations may involve government anchoring initial participation while enabling private extension. The Finnish Maritime Data Space (Section 3.4 of Chapter 3), anchored by the state-owned enterprise Fintraffic, illustrates this logic. In this sector, government operates alongside private participants rather than as sole operator. Several examples across sectors illustrate how this channel operates in practice.

Estonia's X-Road, operational since 2001 and now governed by the Nordic Institute for Interoperability Solutions (comprising Estonia, Finland and Iceland), provides a longer-running example. Serving approximately 52 000 organisations with nearly 3 000 e-services and processing more than 1 billion data transactions per year, X-Road is implemented in over 20 countries (e-Estonia, n.d.^[37]). It functions as a data space under the broad definition proposed in Section 2.3 of Chapter 2: a governance framework, supported by enabling technical and organisational services, that facilitates trusted and interoperable data sharing among participants. Its evolution is now converging with the narrow definition: the X-Road 8.0 "Spaceship" initiative is aligning its trust framework with data space protocol stacks, including Gaia-X compatibility (NIIS, 2025^[38]). X-Road illustrates how long-running government interoperability infrastructure can provide the institutional foundation on which data space capabilities are incrementally layered, a trajectory that inverts the more common pattern of building data space architecture from scratch. It also shows how integration with legacy government systems can proceed incrementally rather than through wholesale replacement.

The EHDS regulation, which entered into force in March 2025, takes a different approach by legislatively mandating government bodies as data space operators. Member States must designate national digital health authorities (for primary use in healthcare delivery) and health data access bodies (for secondary use in research), with the European Commission developing and maintaining central cross-border infrastructure. This is the first EU regulation that explicitly requires establishment of government-operated data space components. It creates a distinct governance model from the consortium-led approach used for most other EU sectoral data spaces.

The EU Public Procurement Data Space (PPDS), launched in September 2024, is among the most documented cases of government as direct data space operator. Managed by the European Commission (DG GROW), the PPDS connects national procurement databases with EU-level procurement data from Tenders Electronic Daily through a federated data layer built on the eProcurement Ontology as common semantic framework. Over 250 000 public authorities engage in public procurement across the EU, spending an estimated EUR 2.5 trillion annually. This represents around 15% of gross domestic product, making it one of the largest potential data spaces by transaction volume (European Commission, n.d.^[39]). EU countries can connect their national databases on a voluntary basis; an OECD-implemented project funded by the EU Technical Support Instrument is supporting a cohort of Member States in connecting to the PPDS (OECD, n.d.^[40]). The PPDS is notable not only as a government-operated data space; the government's own procurement activity also generates the data that populate the infrastructure. Consequently, the government is simultaneously operator, data provider and primary user.

Box 4.2. Inter-governmental (G2G) data-sharing frameworks in the public sector

Inter-governmental data-sharing frameworks are increasingly being developed to enable co-ordinated use of data across government entities. These aim to overcome government data silos and promote more effective data use within the government. Such approaches support interoperability, accountability and data stewardship, while reducing unnecessary centralisation of sensitive information. In so doing, they enable trusted and scalable data sharing across public authorities.

From a conceptual perspective, whether such arrangements meet the narrow definition of data spaces (Section 2.3 of Chapter 2) depends on their governance design. Public authorities are typically legally distinct organisations with separate mandates and accountability structures, and separate data controllers. An inter-governmental framework in which each authority retains custody and control over its data under a shared co-ordination framework can therefore satisfy the federated governance feature of the narrow definition. Conversely, only the broad definition applies where a single lead agency concentrates governance authority and operates a central repository, or where usage conditions are fixed centrally rather than defined by the providing authorities. In either configuration, such frameworks support scalable, co-ordinated and trusted data sharing and use in the public sector in an effective manner.

Israel's inter-ministerial framework in the construction sector illustrates the federated configuration: it aggregates and shares planning and infrastructure data across government entities while maintaining data management within the respective authorities. This approach combines domain-level data control with cross-government co-ordination to improve efficiency and planning outcomes (Government of Israel, 2021^[41]; Government of Israel, 2025^[42]).

Collective and infrastructural value can also characterise private or mixed data spaces. However, purely public sector data-sharing frameworks differ in that such value cannot be captured through market-based mechanisms. As a result, their sustainability does not rely on revenue generation or cost recovery through user participation, but rather on public funding justified by the public value they generate. In these contexts, data-sharing initiatives are assessed primarily in terms of their contribution to public value, such as improved administrative efficiency, policy coherence and enhanced service delivery. For their part, considerations of financial viability remain relevant insofar as they inform the effective and proportionate use of public resources (OECD, 2019^[43]).

The funding implication is that purely public sector data infrastructure, including maintenance of the data themselves, is sustained as a permanent utility through recurrent public budgets, in the same way as other standing government infrastructure. This contrasts with time-limited project funding whose co-ordination limits Section 4.2 discusses. Estonia's X-Road, continuously operated and publicly funded for more than two decades, illustrates the durability of this model.

In the United States, the Genesis Mission – established by Executive Order in November 2025 (The White House, 2025^[44]) – illustrates a government-operator approach focussed on AI-relevant proprietary data rather than public sector transactions. The US Department of Energy is building a federated platform connecting the 17 national laboratories. In so doing, it aims to combine federal scientific datasets with industry and academic contributions through public-private partnerships to train AI for priority innovation sectors. As its distinguishing feature, the programme explicitly raises the question of intellectual property attribution for contributed data, a challenge that has constrained comparable efforts worldwide (Section 3.3 of Chapter 3). Implementation is at an early stage, but the programme provides another model in which government acts as anchor operator for AI-oriented data sharing without adopting the data spaces label.

Government procurement can also serve as a demand-side driver for data space adoption, complementing the supply-side instruments discussed in Section 4.2.1. The Interoperable Europe Act (European Union, 2024^[45]), which entered into force in April 2024, mandates interoperability assessments for EU and national digital public services, effectively requiring the infrastructure that underpins data spaces without explicitly naming them. The eFTI Regulation on electronic freight transport information similarly requires Member State authorities to accept information shared electronically by operators via certified platforms by July 2027, creating de facto demand for data-space type infrastructure in logistics. Where governments require interoperable, standards-based data exchange in their own procurement and operations, they create market demand for data space components that private sector providers can then extend to other contexts.

More generally, existing data-sharing initiatives offer infrastructure and participation that data spaces can build on. Some legislation or sectoral arrangements already require or organise the publication or exchange of data. By making the associated infrastructure compatible with data space protocols, this approach could add usage control, cross-border federation and mechanisms for value creation, without requiring new legislation. The evolution of Estonia's X-Road towards data space compatibility, discussed above, illustrates this trajectory for government interoperability infrastructure.

References

- Bisière, C. et al. (2025), “The economics of data spaces”, *Policy Paper*, SE Digital Center and Toulouse School of Economics, France, https://www.tse-fr.eu/sites/default/files/TSE/documents/DigitalCenter/policy_paper/the_economics_of_data_spaces_july_2025_policy-paper.pdf. [12]
- Bpifrance (2023), “Appel au projets: Espace de données”, <https://www.bpifrance.fr/nos-appels-a-projets-concours/appel-a-projets-espace-de-donnees> (accessed on 14 April 2026). [20]
- Bundesnetzagentur (2022), “Funding competition: Innovative and practical applications and data spaces in the Gaia-X digital ecosystem”, <https://www.bundesnetzagentur.de/EN/Areas/Digitalisation/Data/GaiaX/start.html> (accessed on 14 April 2026). [19]
- CoE-DSC (2024), *Landscape Scan: Dutch Data Sharing Initiatives & Service Providers, initial trial version*, <https://coe-dsc.nl/wp-content/uploads/2025/01/Dutch-Data-Sharing-Landscape-Scan-2024-report.pdf>. [16]
- Digital Research Alliance of Canada (2025), “National data spaces pilot funding program: Call for proposals”, <https://www.alliancecan.ca/en/opportunities/national-data-spaces> (accessed on 16 April 2026). [21]
- DSSC (2025), “Participant adoption and sustainability in data spaces”, <https://dssc.eu/publications> (accessed on 16 April 2026). [18]
- DSSC (2024), “Generative AI and data spaces”, <https://dssc.eu/space/Generative+AI+and+Data+Spaces> (accessed on 14 April 2026). [17]
- e-Estonia (n.d.), “X-Road”, <https://e-estonia.com/solutions/interoperability-services/x-road/> (accessed on 16 April 2026). [37]
- European Commission (2025), “Data union strategy: Unlocking data for AI”, <https://digital-strategy.ec.europa.eu/en/library/data-union-strategy-unlocking-data-ai> (accessed on 6 March 2026). [9]
- European Commission (2025), *Proposal for a Regulation amending Regulations (EU) 2022/868, (EU) 2023/2854 and Directive (EU) 2019/1024 (Digital Legislation Omnibus)*, European Commission, Brussels. [8]
- European Commission (n.d.), “Simpl: Cloud-to-edge federations empowering EU data spaces”, <https://digital-strategy.ec.europa.eu/en/policies/simpl> (accessed on 14 July 2026). [47]
- European Commission (n.d.), “The Public Procurement Data Space (PPDS) – Internal market, industry, entrepreneurship and SMEs”, https://single-market-economy.ec.europa.eu/single-market/public-procurement/digital-procurement/public-procurement-data-space-ppds_en (accessed on 31 March 2026). [39]
- European Union (2025), *Regulation (EU) 2025/327 on the European Health Data Space*, European Union, Brussels. [1]

- European Union (2024), *Artificial Intelligence Act European Parliament legislative resolution of 13 March 2024 on the proposal for a regulation of the European Parliament and of the Council on laying down harmonised rules on Artificial Intelligence (Artificial Intelligence Act) and amending certain Union Legislative Acts (COM(2021)0206 – C9-0146/2021 – 2021/0106(COD))*, European Union, Brussels, https://www.europarl.europa.eu/doceo/document/TA-9-2024-0138_EN.pdf. [6]
- European Union (2024), *Regulation (EU) 2024/1183 amending Regulation (EU) No 910/2014 as regards establishing the European Digital Identity Framework*, European Union, Brussels, <https://eur-lex.europa.eu/eli/reg/2024/1183/oj>. [36]
- European Union (2024), *Regulation (EU) 2024/1781 establishing a framework for setting ecodesign requirements for sustainable products (Ecodesign for Sustainable Products Regulation)*, European Union, Brussels. [13]
- European Union (2024), *Regulation (EU) 2024/903 laying down measures for a high level of public sector interoperability across the Union (Interoperable Europe Act)*, European Union, Brussels. [45]
- European Union (2023), *Regulation (EU) 2023/1542 concerning batteries and waste batteries*, European Union, Brussels. [3]
- European Union (2023), *Regulation (EU) 2023/2854 on harmonised rules on fair access to and use of data (Data Act)*, European Union, Brussels. [7]
- European Union (2022), *Regulation (EU) 2022/868 of the European Parliament and of the Council of 30 May 2022 on European data governance and amending Regulation (EU) 2018/1724 (Data Governance Act)*, European Union, Brussels, <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32022R0868>. [5]
- European Union (2019), *Directive (EU) 2019/1024 of the European Parliament and of the Council of 20 June 2019 on open data and the re-use of public sector information*, European Union, Brussels, <https://eur-lex.europa.eu/legal-content/EN/TXT/?qid=1561563110433&uri=CELEX:32019L1024>. [10]
- European Union (2018), *Regulation (EU) 2018/1807 on a framework for the free flow of non-personal data in the European Union*, European Union, Brussels. [11]
- Government of Israel (2025), *Infrastructure for Growth 2026*, Government of Israel, https://www.gov.il/BlobFolder/generalpage/plan-infrastructure-2025/he/files_Infrastructure240326.pdf. [42]
- Government of Israel (2021), “Aggregation and sharing of information for planning, implementation, and improvement of regulation in the construction sector”, Government Decision No. 854 (1 March 2021), Government of Israel, https://www.gov.il/he/pages/dec854_2021. [41]
- HHS (2026), “TEFCA – America’s national interoperability network reaches nearly 500 million health records exchanged”, <https://www.hhs.gov/press-room/tefca-americas-national-interoperability-network-reaches-nearly-500-million-health-records-exchanged.html> (accessed on 7 April 2026). [4]
- Holmström, B. (1979), “Moral hazard and observability”, *The Bell Journal of Economics*, Vol. 10/1, p. 74, <https://doi.org/10.2307/3003320>. [46]

- Holmström, B. and P. Milgrom (1991), “Multitask principal-agent analyses: Incentive contracts, asset ownership”, *Journal of Law, Economics, & Organization*, Vol. 7/special issue, pp. 24-52, <https://www.sfu.ca/~allen/HolmstromMilgrom.pdf>. [25]
- IPA (2026), “Open Data Spaces Reference Architecture Model (ODS-RAM)”, <https://open-dataspaces.gitbook.io/ods-docs/ods-ram/v2> (accessed on 1 April 2026). [32]
- IPA (2025), “Agreement to jointly promote the dataspace technical concept as ‘open data spaces’”, <https://www.ipa.go.jp/en/pressrelease/press20251015.html> (accessed on 9 April 2026). [30]
- IPA (2024), “Data Spaces Academy”, <https://www.ipa.go.jp/en/digital/data/data-spaces-academy.html> (accessed on 28 April 2026). [29]
- Lerner, J. (2009), *Boulevard of Broken Dreams: Why Public Efforts to Boost Entrepreneurship and Venture Capital Have Failed — and What to Do About It*, Princeton University Press. [26]
- Mertens, C. et al. (2024), “Data spaces business models”, *Position Paper*, Zenodo (CERN European Organization for Nuclear Research), <https://doi.org/10.5281/zenodo.14101303>. [27]
- METI (2025), “Outstanding initiatives to promote industrial data sharing selected under the project certification of Ouranos ecosystem”, https://www.meti.go.jp/english/press/2025/0509_002.html (accessed on 22 April 2026). [34]
- METI/IPA (2025), *Open Data Spaces Reference Architecture Model (ODS-RAM)*, Ministry of Economy, Trade and Industry (METI) and the Information-Technology Promotion Agency, Government of Japan. [31]
- Moonen, N. et al. (2025), *Sustainable Revenue Models for Data Sharing Initiatives*, Jheronimus Academy of Data, Tilburg University, The Netherlands, https://www.government.nl/site/binaries/site-content/collections/documents/2025/04/22/sustainable-revenue-models-for-data-sharing-initiatives/Report_Sustainable_Revenue_Models_for_Data_Sharing_Initiatives.pdf. [24]
- National Data Administration (2024), *Trusted Data Space Development Action Plan (2024-2028)*, National Data Administration, People’s Republic of China. [22]
- NEO-UFRGS (2024), *Condições estruturantes para o desenvolvimento de data spaces no Brasil: Uma visão de especialistas [Green Paper: Structuring data spaces in Brazil]*, Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ) GmbH on behalf of the German Federal Ministry for Digital and Transport (BMDV), <https://digital-dialogues.net/en/news-details/publication-green-paper-on-structuring-data-spaces>. [35]
- NIIS (2025), “X-Road 8 Spaceship”, <https://x-road.global/spaceship> (accessed on 22 April 2026). [38]
- NITI Aayog (2020), *Data Empowerment and Protection Architecture: Draft for Discussion*, NITI Aayog, Government of India, <https://www.niti.gov.in/sites/default/files/2023-03/Data-Empowerment-and-Protection-Architecture-A-Secure-Consent-Based.pdf>. [2]
- OECD (2019), *Public Value in Public Service Transformation: Working with Change*, OECD Publishing, Paris, <https://doi.org/10.1787/47c17892-en>. [43]

- OECD (2017), *Recommendation of the Council on Health Data Governance*, OECD/LEGAL/0433, <https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0433> (accessed on 10 March 2026). [14]
- OECD (n.d.), “Leveraging the EU Public Procurement Data Space (PPDS)”, <https://www.oecd.org/en/about/programmes/sg-reform/leveraging-the-eu-public-procurement-data-space-ppds.html> (accessed on 10 April 2026). [40]
- Otto, B. and H. Tardieu (2025), *Data Sharing: Closing the Gap in the Draghi Report’s Call for AI Competitiveness in Europe*, Le Grand Continent. [23]
- Red.es (2025), “Kit Espacios de Datos”, (Kit Data Spaces), <https://www.red.es/es/iniciativas/proyectos/kit-espacios-de-datos> (accessed on 25 April 2026). [28]
- Sitra (2024), “Gaia-X Finland builds one of Europe’s leading data sharing business networks in Finland”, <https://www.sitra.fi/en/articles/gaia-x-finland-builds-one-of-europes-leading-data-sharing-business-networks-in-finland/> (accessed on 15 April 2026). [15]
- The White House (2025), “Launching the Genesis mission”, <https://www.whitehouse.gov/presidential-actions/2025/11/launching-the-genesis-mission/> (accessed on 28 April 2026). [44]
- Tsuda, M. (2026), *Why Open Dataspaces: Design Philosophy and Architectural Paradigm*, Information-technology Promotion Agency, Tokyo, https://www.ipa.go.jp/en/digital/architecture-guidelines/individual-link/p1o1lf000001xv4n-att/WhyOpenDataspaces_en.pdf. [33]

Endnotes

¹ Another example of EU funding includes *Simpl*, the open-source smart middleware procured by the European Commission to provide common infrastructure for the common European data spaces and cloud-to-edge federations (European Commission, n.d.^[47]).

² “Moral hazard” is used here in its principal-agent sense (Holmström, 1979^[46]): The funder (principal) cannot directly observe whether recipients (agents) are allocating effort towards operational viability or towards satisfying formal milestones. Holmström and Milgrom (1991^[25]) formalise this as a multi-task principal-agent problem: when some tasks are more measurable than others, incentives tied to the measurable task crowd out effort on the unmeasurable one. In this context, compliance is measurable; progress towards operational viability is not.

5

Conclusions and policy considerations

This chapter draws together the findings of the report and sets out eight policy considerations for governments seeking to support viable data spaces. These concern targeting public investment where a genuine co-ordination problem exists, distinguishing funding for the technical foundation from funding for ecosystems, addressing the demand-side funding gap, supporting business models around infrastructural value, promoting trust-enhancing technologies, using certification and government anchoring to build institutional trust, recognising AI-driven data demand as a pathway to viability, and enhancing measurement. The considerations support OECD work on data free flow with trust (DFFT) and the OECD Recommendation on Enhancing Access to and Sharing of Data (EASD).

Governments in the European Union (EU) alone have invested an estimated EUR 1.7 billion in a data space ecosystem whose value they cannot yet measure. Only one in nine entries has reached operational status. Moreover, not a single initiative has reached the scaling stage that the definitions of the International Data Spaces Association itself associate with viability. Finally, the evidence base on what has been built remains thin (Box 1.1 of Chapter 1). The technology has been demonstrated to work, but evidence is lacking on whether the investment is delivering the promised returns.

Government intervention may be justified where genuine co-ordination failure exists that alternative approaches leave structurally unresolved. In domains such as supply chain traceability, health data exchange, multimodal mobility and energy system integration, the parties that need to share data are often competitors or jurisdictional rivals unwilling to cede control to a single commercial intermediary. Data spaces address this class of problem through distributed governance, open interoperability and provider-defined usage control (Section 3.3 of Chapter 3). Regulatory requirements across multiple jurisdictions amplify the urgency of solving these co-ordination problems (Section 4.1 of Chapter 4).

The policy considerations below address how governments can sharpen their approach to targeting, funding, measuring and delimiting data space initiatives.

5.1. Targeting investment where the structural condition holds

Four boundary conditions define the class of problem for which data spaces have a comparative advantage over alternative mechanisms: (i) three or more parties; (ii) competitive or jurisdictional relationships that preclude trusting a single commercial intermediary; (iii) non-excludable collective value that bilateral contracts cannot capture; and (iv) ongoing data flows requiring continuous governance rather than one-off transactions (Section 3.2 of Chapter 3). These conditions are particularly pertinent in cross-border settings, where regulatory divergence compounds the co-ordination problem and no single intermediary can easily claim cross-jurisdictional authority (Section 4.3 of Chapter 4). Where these conditions are absent, alternative approaches may suffice and may not justify the co-ordination costs that data spaces entail (Table 3.1 of Chapter 3).

The presence of the four boundary conditions is necessary but not sufficient, however. Sectoral conversion rates (Figure 3.1 of Chapter 3) suggest that pre-existing digital infrastructure, identifiable anchor firms and clear value propositions among early participants are preconditions for data space viability, irrespective of initiative volume. A two-stage assessment could help direct public resources more effectively towards use cases where data spaces are both structurally necessary and practically viable. The first stage would examine whether the co-ordination problem meets the boundary conditions. The second would consider whether the sectoral preconditions are in place.

5.2. Distinguishing between funding for the technical foundation and for ecosystems

Two distinct purposes of public funding merit separation in policy design (Section 4.2.1 of Chapter 4). Investment in information models, protocols and interoperability frameworks constitutes pre-competitive research and development (R&D) with broad spillover benefits. For its part, investment in the creation of actual data space ecosystems is more conditional and depends on whether a specific collective-action problem has been identified that no market actor has the incentive or legitimacy to solve. Most funding programmes were not explicitly designed around this distinction or the bootstrapping rationale identified in Section 3.4 of Chapter 3. The stated rationale centres on competitiveness, digital sovereignty and, in the case of the European Union, the single market for data, rather than on correcting a formally identified collective-action failure (Section 4.2.1 of Chapter 4).

Where funding instruments reward observable compliance (consortium formation, architectural alignment, dissemination) rather than operational viability and participant willingness to pay, rational actors tend to optimise for the former (Lerner, 2009^[1]). Furthermore, long-term public funding may also create dependency effects that prevent data space stakeholders from developing viable revenue models, as discussed at the OECD Expert Workshop (Annex A). Where governments fund ecosystem creation rather than pre-competitive R&D, success metrics calibrated to operational outcomes (participant growth, data flows, revenue readiness) rather than architectural compliance may better align incentives with viability.

5.3. Addressing the demand-side funding gap

The dominant policy response has been supply-side funding directed at shared infrastructure. Far less common is funding that addresses the onboarding costs that individual participants, particularly small and medium-sized enterprises (SMEs), face when joining a data space (Section 4.2.2 of Chapter 4). The bootstrapping problem operates at two levels. The collective investment in shared infrastructure must precede the network effects that justify it. Moreover, each participant's private onboarding investment must be justified by expected returns that may not materialise until the ecosystem reaches sufficient scale (Section 3.4 of Chapter 3). Addressing only the first level while leaving the second unfunded may help explain why data spaces that receive substantial infrastructure investment nonetheless fail to attract sufficient participation.

Three demand-side approaches have emerged across jurisdictions: direct subsidy, public investment and regulatory design. Direct subsidy for participant onboarding costs is illustrated by Spain's Data Spaces Kit, which subsidises integration of individual firms into sectoral data spaces on a non-competitive basis. Public investment in capacity building and data maturity are features of Japan's educational programme targeting businesses and SMEs (Information-technology Promotion Agency). Finally, regulatory design can create commercially viable intermediary categories rather than relying on public subsidy. In Data Empowerment and Protection Architecture (DEPA) in India, for example, regulated consent managers generate revenue from facilitating data exchange. Given these multiple approaches, governments may wish to complement supply-side infrastructure funding with instruments that reduce participant onboarding costs. They could do this through direct subsidy, capacity building or regulatory frameworks that make the intermediary role commercially viable.

5.4. Supporting business models around infrastructural value

The value that data spaces produce is predominantly collective and infrastructural, exhibiting properties that Frischmann (2012^[2]) associates with shared resources whose social benefits exceed what pricing mechanisms can capture. Business models premised on data monetisation assume bilateral, transactional value. Where value is collective and non-market, these models risk being structurally misaligned with the value proposition they are meant to serve. Empirical evidence from European data-sharing ecosystems is consistent with this misalignment: in a qualitative study across mobility, aerospace, energy, agriculture and other sectors, 8 of 20 initiatives framed themselves as data marketplaces. However, their orchestrators acknowledged that the initial goal had been deeper collaboration rather than transactional exchange (Eustache, Brousseau and Toledano, 2025^[3]).

Viable models observed in operational data spaces resemble cost-sharing arrangements in which participants jointly fund the infrastructure and share in system-level benefits (the Finnish Maritime Data Space); membership-based access to co-ordination infrastructure (BoostAeroSpace); or hybrid models combining public co-investment with participant cost recovery. Prioritising use-case value over infrastructure development is associated with stronger prospects for financial viability (DSSC, 2026^[4]), a finding confirmed at the OECD Expert Workshop (Annex A).

That many data spaces depend on public funding does not in itself demonstrate that they are public infrastructure. It may instead indicate that viable business models have not yet been found. In particular, funding logic may need to evolve over the ecosystem life cycle, shifting from early-stage subsidies and neutral cost sharing during emergence towards usage-based or performance-linked mechanisms as the ecosystem matures (Eustache, Brousseau and Toledano, 2025^[3]). Governments may wish to distinguish between use cases where data monetisation is viable and where value is collective and infrastructural, and to calibrate funding expectations accordingly.

5.5. Supporting adoption of trust-enhancing technologies and their integration by design

Provider-defined usage control is a major capability that distinguishes data space interoperability from both commercial platform interoperability and traditional industry consortium data exchange (Section 3.3.3 of Chapter 3). Without credible guarantees that shared data will be used only as agreed, rational actors withhold participation regardless of the potential collective benefit. Two complementary approaches are maturing in parallel: data usage control (e.g. built on Open Digital Rights Language and dataspace protocols); and privacy-enhancing technologies (PETs) such as federated learning, secure multi-party computation and trusted execution environments. Neither alone is sufficient.

Sustained support for the maturation of both tracks, and for the interoperability layer that allows them to be combined within a single data space deployment, would strengthen the trust foundations on which participation depends. This observation is based on the support of the data free flow with trust (DFFT) Expert Community on OECD work related to PETs, as well as a complementary OECD project on “Trust by Design” in cross-border supply chains [DSTI/DPC/DGP(2026)5]. Governments are already recognising the importance of PETs for data sharing, including in cross-border contexts. The G7 Industry, Digital and Technology Ministerial Declaration of December 2025 commits to “advanc[ing] DFFT through the promotion of privacy-enhancing technologies (PETs)” and calls for “initiatives that support greater understanding of PETs, such as the OECD Global Repository on PETs” (G7, 2025).

5.6. Adopting certification and government anchoring to build institutional trust

Every data space that has reached operational viability features a credible neutral anchor entity. This could be a state-owned enterprise, a non-profit industry consortium or a large firm whose market position makes neutrality credible (Section 3.3.1 of Chapter 3). In some sectors, no natural anchor exists, whether because of participant fragmentation, heterogeneity or jurisdictional dispersal. In such cases, governments may wish to consider designating public bodies or state-owned enterprises to serve as institutional anchors (Section 4.5 of Chapter 4).

Estonia’s X-Road and the European Health Data Space regulation illustrate two models of institutional anchors: long-running government interoperability infrastructure on which data space capabilities can be layered; and legislative mandates that designate government bodies (digital health authorities, health data access bodies) as the institutional infrastructure for data sharing. Deploying data spaces for the government’s own operations – from inter-governmental (G2G) configurations to arrangements extending to businesses and citizens (G2G2B) – reinforces this channel by leading by example. To that end, it demonstrates that the architecture works at scale, legitimises the underlying standards; and lowers the perceived adoption risk for private participants (Section 4.5 of Chapter 4).

Certification is emerging as a complementary instrument. Japan’s “public interest digital platform operator” certification (Section 4.4 of Chapter 4), under which the Automotive and Battery Traceability Centre has been accredited (Box 4.1 of Chapter 4), assesses operators against substantive criteria for reliability and

public benefit. This generates a trust signal that may help address the bootstrapping problem. Governments developing data space frameworks may wish to consider how certification or registration mechanisms can provide the institutional trust that the market may not be able to generate on its own.

5.7. Recognising AI-driven data demand as a pathway to viability

AI is emerging as a catalyst for data space viability. The demand for high-quality, domain-specific training data creates concrete business incentives for governed data sharing that previous waves of data space formation lacked. No single actor possesses the full data resources needed to train competitive AI models in industrial sectors, making collaborative data ecosystems an economic necessity rather than a policy aspiration. The EU Data Union Strategy's repositioning of data spaces as AI supply chain infrastructure, including the introduction of "data labs" to link data spaces to AI ecosystems (Section 4.2.1 of Chapter 4), signals institutional recognition of this symbiosis. Similar institutional recognition is visible in Japan, where the ODS-RAM v2 design philosophy situates the architecture explicitly in relation to agentic AI (Tsuda, 2026^[5]).

Government, including through funding programmes and viability assessments, could encourage data space initiatives to articulate their value proposition in terms of the AI use cases they enable, where these exist. Whether AI-driven demand will prove transformative depends on whether data spaces can deliver governed data access at the speed and scale that AI development requires. The current ecosystem still needs to demonstrate this condition via operational use cases.

5.8. Enhancing measurement

Despite an estimated EUR 1.7 billion in EU public investment alone, few initiatives collect baseline data on the co-ordination problem they aim to solve; fewer still measure outcomes against that baseline in terms that allow comparison across sectors or with alternative data-sharing mechanisms. Governments may wish to condition future public funding on the systematic collection of baseline and outcome data sufficient to evaluate whether the initiative is delivering value that alternative mechanisms could not. The geographic dimension compounds the problem. The evidence base is overwhelmingly European; non-European architectures operating on comparable principles (e.g. Open Data Spaces but also the Trusted Exchange Framework and Common Agreement and DEPA) are absent or underrepresented in the primary quantitative source for this paper. Extending systematic data collection on federated data-sharing initiatives across OECD Members and partners is a precondition for evidence-based policy guidance on an area that is genuinely international in scope.

References

- DSSC (2026), “Data Spaces Blueprint v3.0”, [4]
<https://dssc.eu/space/BVE/357074549/Data+Spaces+Blueprint+v3.0> (accessed on 9 April 2026).
- Eustache, L., E. Brousseau and J. Toledano (2025), “The economics of data-sharing: An empirical investigation of data sharing ecosystems”, *Forty-Sixth International Conference on Information Systems (ICIS)*, Nashville, Tennessee, <https://hal.science/hal-05542222v1>. [3]
- Frischmann, B. (2012), *Infrastructure: The Social Value of Shared Resources*, Oxford University Press. [2]
- Lerner, J. (2009), *Boulevard of Broken Dreams: Why Public Efforts to Boost Entrepreneurship and Venture Capital Have Failed — and What to Do About It*, Princeton University Press. [1]
- Tsuda, M. (2026), *Why Open Dataspaces: Design Philosophy and Architectural Paradigm*, Information-technology Promotion Agency, Tokyo, [5]
https://www.ipa.go.jp/en/digital/architecture-guidelines/individual-link/p1o1lf000001xv4n-att/WhyOpenDataspaces_en.pdf.

Annex A. OECD Workshop on “Assessing the Sustainability of Business Models in Data Spaces and the Role of Public Policies”

Context

The OECD Expert Workshop on “Assessing the sustainability of business models in data spaces and the role of public policies” was held on 11 March 2025 at the Data Spaces Symposium 2025 in Warsaw, Poland. The workshop brought together practitioners, policymakers and industry experts from Europe, Japan and North America to critically examine the financial sustainability of data space business models on the basis of concrete use cases, both successful and failed. For the agenda, see Data Spaces Symposium (2025^[1]).

Business models

No single business model for data spaces exists. The value proposition, cost structure and revenue model depend on the sector, the use case, the participants involved and the regulatory context. A central finding was that value is created at the level of individual use cases, not at the level of the infrastructure itself; prioritising use-case value over infrastructure development was consistently associated with stronger prospects for sustainability.

An imbalance frequently arises between those who bear the costs of contributing data and those who capture the value of using them; this was identified as a persistent challenge. Each stakeholder tends to understand its own business model but often lacks a view of the overarching business value. At least one orchestrator or champion who understands both the overall value proposition and the addressable market is therefore necessary. However, distributing the benefits created by a data space requires complex negotiation among members, including orchestrators who may also have their own business interests.

Regulatory compliance is a primary driver of participation, but building a business model primarily on compliance carries risks. In markets dominated by small and medium-sized enterprises (SMEs), regulations may not serve as an effective incentive: SMEs often need to digitise their operations first, and additional requirements only increase costs and reduce their capacity to innovate with shared data.

Data spaces face a classic chicken-and-egg bootstrapping problem. Strategies discussed included attracting participants who act on both sides of the market; leveraging existing platforms; deploying high-value initial use cases (“seeding”); securing “marquee users”; and cross-subsidisation during early adoption. One approach involved keeping the data space operator deliberately compact, connecting “collaborative services” (shared infrastructure guided by public guidelines) with “competitive services” offered by third-party providers through open application programming interfaces. A recurring challenge is “reinventing the wheel”: consortia whose members vary across initiatives spend excessive time debating technical components from scratch.

Use-case insights

Automotive and battery traceability (Japan). A data space established by Japanese automotive manufacturers to compile carbon footprint data for EU Battery Regulation compliance was presented as a model of sequenced expansion: starting with battery carbon footprint traceability (“collaborative services”), then expanding to other vehicle components and eventually to “competitive services” such as battery passports and supply chain resilience. Its revenue model combines membership fees with transaction-based data-use fees. The operator adopted an existing Open Data Spaces Reference Architecture Model (former Ouranos Ecosystem) for technical standards and obtained accreditation as a “public interest digital platform operator.”

Maritime data sharing (Finland). The Finnish Maritime Data Space was highlighted as an example where the business model was designed before the technology. Five sequential steps were identified: (i) identify ecosystem champions from shipping companies, charterers, ports, cargo owners and relevant public entities; (ii) develop systemic service design by mapping stakeholders’ challenges, roles, data holdings and contractual frameworks; (iii) map the real-life contractual model and value model; (iv) build the business model, including a revenue-sharing arrangement among participants; and (v) then enable the solution through data space technology. This process took eight months. Key factors for industry buy-in were estimated savings of EUR 45 million in fuel and EU Emissions Trading System costs per year (2022 baseline for Finland) and CO₂ emission reductions exceeding 24%.

Agricultural data (France). The bankruptcy of a prominent agricultural data space operator in November 2024 served as a cautionary case. The operator had focussed on building technical infrastructure without demonstrated use-case value sufficient to sustain participation. The agricultural sector’s fragmentation into many SMEs made it inherently harder to organise into a data-sharing ecosystem than sectors with concentrated global companies sharing common supply chains.

Policy opportunities and challenges

Participants distinguished between two purposes of public funding: the development of schemas, protocols and interoperability frameworks (pre-competitive research and development), and the creation of actual data space ecosystems, where the case for public funding is more conditional. Not all data spaces can be considered public goods; where many depend on public funding, this may indicate that viable business models have not yet been found rather than that they constitute public infrastructure. Some participants cautioned that long-term public funding may create dependency effects that prevent data space stakeholders from being efficient and innovative.

Japan’s certification system for “Public Interest Digital Platform Operators,” administered by the Information-technology Promotion Agency (IPA), was discussed as a potential model for other jurisdictions seeking to distinguish data spaces that serve a public interest from those that are purely commercial. IPA has also invested in capacity building (data space guides, handbooks for Chief Data Officers, data maturity frameworks). The rationale is that businesses first need to understand what data spaces are before they can participate, given that low awareness, limited skills and low data maturity were identified as significant barriers among SMEs.

Participants drew an analogy with the Internet, arguing that standardisation discussions should focus on the protocol layer rather than on how protocols are implemented or used in business. This implies a “coopetition” model in which co-operation on protocols coexists with competition on services. Concern was expressed about a trend of de-globalisation, a shrinking willingness to share data across jurisdictions. Differences in regulations between countries were identified as having a potentially negative impact on global manufacturing industries. Participants observed that intra-organisational data ecosystems in the United States appear to exist in business but are not named as data spaces. This suggests the concept may need to be framed differently outside the European context.

References

Data Spaces Symposium (2025), “Memories of Data Spaces Symposium 2025”,
<https://www.data-spaces-symposium.eu/program-2025> (accessed on 1 March 2026). [1]

Data Spaces between Co-ordination Challenges, Financial Viability and Public Policy

Sharing data across organisations and jurisdictions generates economic and social value, increasingly so as progress in artificial intelligence depends on access to high-quality, domain-specific data. Data spaces promise trusted multi-party data sharing under federated governance, yet, despite significant public investment, few have reached operational and financial viability. Drawing on the analysis of more than 200 initiatives, including case studies of successes and failures, and an OECD Expert Workshop, this report examines where data spaces provide value that alternative data-sharing mechanisms cannot. In so doing, it illustrates how data spaces best suit settings with three or more parties, limited trust in any single intermediary, collective value creation and ongoing governance needs. Governance design, interoperability and trust-enhancing mechanisms underpin this value, but many face a bootstrapping problem, which eight policy considerations help address. The report supports OECD work on data free flow with trust (DFFT) and the OECD Recommendation on Enhancing Access to and Sharing of Data.



PRINT ISBN 978-92-64-76919-9
PDF ISBN 978-92-64-95733-6

