INTERNATIONAL DATA SPACES ASSOCIATION

Position Paper | Version 1.0 | July 2026

Data Spaces and AI
Trustworthy Agentic Participation in Data Spaces

- ○ Position Paper of members of the IDS Association
- ○ Position Paper of bodies of the IDS Association
- ● Position Paper of the IDS Association
- ○ White Paper of the IDS Association

Publisher

International Data Spaces Association
Emil-Figge Straße 80
44227 Dortmund
Germany

Editor

Anil Turkmayali, IDSA

Cite as

Turkmayali A., Data Spaces and AI: Trustworthy Agentic Participation in Data Spaces, International Data Spaces Association, July 2026
https://doi.org/10.5281/zenodo.21279055

Authors & Contributors

Reinhold Achatz, IDSA
Daniel Arosa Otero, NTT Data
Arian Artman, NTT Data
Silvia Castellvi, IDSA
Masaru Dobashi, NTT Data
Liu Dong, China Future Internet Engineering Center
Felix Dreher, SICK AG
Roland Fadrany, GAIA-X
Olaf Gerd-Gemein, Orbiter
Haluk İnanmış, FS Software
Petteri Kivimäki, Nordic Institute for Interoperability Solutions
Peter Koen, Microsoft
Gabriella Laatikainen, VTT
Viivi Lähteenoja, 1001 Lakes
Fernando López Aguilar, GMV GmbH
Christoph Mertens, IDSA
Achim Pascal Meyer, Sphin-X
Yasser Mohammad, NEC
Lars Nagel, IDSA
André Nemat, IDSA
Akira Sakaino, NTT DOCOMO BUSINESS
Christoph Schlueter Langdon, Drucker School and Deutsche Telekom
Kosmas Tsiakas, Centre for Research and Technology Hellas
Dimitrios Giakoumis, Centre for Research and Technology Hellas
Marko Turpeinen, 1001 Lakes
Markus Ketterl, msg systems
Sohvi Virkkula, VTT
Giuseppe Zangari, Aruba

Disclaimer

The views expressed in this paper are those of the individual authors and do not necessarily reflect the positions of their organizations.

Contributing organizations

Executive Summary

Two developments are reshaping how organizations work with data, and they are usually discussed separately. Data spaces give independent organizations a way to share data and data-based services under agreed rules, with each participant keeping control of their own assets. Artificial intelligence is being adopted across every sector, but its value depends on access to high-quality data with clear usage rights, reliable provenance and a sound basis for compliance. This paper connects the two, because each supplies what the other needs. In the AI era, the fundamental unit of exchange is progressively shifting from data to governed intelligence.

The relationship runs in both directions. Data spaces give AI a governed, high-quality data foundation. Catalogues make data discoverable across organizational boundaries, federated identity mechanisms extend trust where it cannot otherwise be assumed, and machine-readable policies make usage rights explicit and enforceable. In return, AI gives data spaces the automation they need to operate and scale, by generating metadata, aligning vocabularies, interpreting policies and reducing the manual effort that has kept cross-organizational data sharing slow and costly.

The frontier of this relationship is agentic participation. An AI agent can act through the same interfaces a human participant uses, faster and without interruption: finding data, negotiating terms and invoking services on an organization's behalf. The central argument of this paper is that this can be made trustworthy without inventing new and untested machinery. An agent operates under a delegated identity that binds it to the legally accountable participant it acts for, so every action it takes remains attributable to a real organization and is governed by the same policies, credentials and audit trails that already secure human participation.

This is a position paper, not a legal opinion or a full technical specification. It establishes a shared vocabulary for readers coming from either side of this convergence, sets out today's AI challenges and how the building blocks of a data space address them, details the concrete value AI brings to operating a data space and grounds the discussion in pilots already underway. For policy makers, it shows why data spaces belong within the agenda for AI governance and innovation. For practitioners on either side, it shows where AI creates new requirements and where it removes effort. The technical and governance foundations largely exist and are being standardized; the work now is to bring them together.

Table of Contents

List of Figures

1 Introduction

1.1 Why was this paper written and for whom

This paper was written because two important topics are converging, while the discussions about them still happen largely in separate places.

The first is data spaces. They were developed to enable trusted data sharing between independent parties, providing a decentralized alternative to platform-based data sharing and letting participants keep control over their data while making it available under agreed conditions. They also address practical obstacles such as interoperability and the enforcement of usage rights and policies.

The second is AI. Organizations are investing heavily in AI systems, but that investment runs into recurring obstacles: a lack of high-quality data, unclear usage rights, and weak traceability and compliance. These are a mix of technical and governance problems.

The purpose of this paper is to show that the two are connected. Data spaces can provide the governed, high-quality data foundation that AI needs, and AI can provide the automation and intelligence that data spaces need to scale. It is written for readers working on one side of this relationship who need to understand the other, whether policy makers, data space practitioners or AI practitioners. Section 1.3 sets out what each will take from it.

This is a position paper, not a legal opinion and not a full technical specification. It offers common framing, identifies the key architectural and governance questions and points to concrete patterns and next steps for the IDSA community and related ecosystems.

1.2 Data spaces, AI, and their interplay: definitions and framing

In this paper, data space is understood as a governed, decentralized environment for data sharing between independent participants. It is not a central silo or a single platform that absorbs all data. The data usually remains with the provider or is transferred only under the usage policies the provider defines.

A data space combines technical and organizational elements: catalogues, connectors, identity mechanisms, access and usage policies, contractual rules, semantic models, protocols, governance bodies and trust services. Together, these create a framework in which participants can exchange data or data-based services while keeping control over their own assets and responsibilities.

AI is used here in a broad but practical sense. It covers classical machine learning, statistical data science, knowledge-based methods, semantic technologies, multi-agent systems, generative AI, large language models, and agentic AI. The paper does not limit AI to one model family or to the current attention around generative AI, because different techniques contribute in different ways: some predict or classify, some support search, matching and anomaly detection, and some handle semantic alignment, metadata generation or policy interpretation. Others operate as agents that plan, call tools, negotiate and interact with data space services.

Scope note. In this paper, AI is not used as a synonym for generative AI. The term covers several families of techniques: statistical and machine-learning methods for prediction,

classification, anomaly detection, and quality assessment; semantic and knowledge-based methods for linking concepts, policies, and metadata; generative models for language, code, and content generation; multi-agent systems for coordination and negotiation; and agentic AI systems that can use tools and act toward a goal. The paper is mainly concerned with the last two families when discussing participation in data spaces, but the operational value of AI for data spaces often comes from the full combination.

The term Agentic AI is used more specifically. An agent is a software-based actor that can pursue a goal within defined boundaries, select actions, use tools, interact with catalogues or services, and possibly negotiate with other participants or agents. In a data space, such an agent must not be treated as an anonymous technical component: its actions must be linked to a participant, a purpose, a scope of authority, and an accountable legal entity.

The interplay between data spaces and AI runs in two directions. In data spaces for AI, data spaces give AI systems access to governed, discoverable data that is usable under explicit conditions, which matters for training, validation and operational decision support and brings better provenance, clearer permissions and stronger accountability. On the other hand, in AI for data spaces, AI supports the creation, operation, and scaling of data spaces: creating and enriching metadata, improving catalogue search, aligning vocabularies, detecting quality issues, translating legal and organizational terms into machine-readable policies, supporting compliance checks, monitoring data flows and assisting participants during onboarding and negotiation.

1.3 How to read this paper

Policy makers and decision-makers should read this paper as an argument for runtime governance. The question is not only whether an AI system is compliant at design time. It is whether every data access, knowledge retrieval, tool call, transformation and publication decision is bound to an accountable participant, an explicit purpose, a policy decision and auditable evidence. Data spaces provide the operating environment in which this holds across organizational boundaries.

Data space practitioners new to AI should read this paper as a preparation guide. The data space primitives already exist: participants, credentials, catalogues, data products, connectors, usage policies, data contracts, provenance, clearing, and certification. The change AI introduces is that some participants will now be autonomous agents rather than only humans or conventional applications. The governance model can stay stable if agent identity, scope, tool access and audit are made explicit.

AI practitioners new to data spaces should read this paper as a caution against direct, ungoverned tool access. A model-connected API is not automatically trustworthy. A Retrieval-Augmented Generation (RAG) pipeline is not automatically compliant. A tool-calling agent is not automatically authorized. Data spaces supply the missing layer: identity, contract, usage control, semantic interoperability, and provenance. The model can propose and orchestrate, but the data space decides and records what is allowed.

2 Data Spaces and AI: Why do they belong together?

Training and operating AI models require access to large volumes of diverse, high-quality data. The more specialized the application, the more critical the quality and provenance of that data become. Much of the most valuable data sits inside organizations that will not share it without legal clarity about permitted use cases, liability, and the boundaries of data sovereignty. Without a governed access layer, this data remains locked.

Data spaces are designed to address this gap. Unlike centralized platforms, they enable providers and consumers to negotiate access terms directly, for each dataset and each intended use. Rights holders can calibrate those terms to the sensitivity and value of each asset, rather than choosing between full disclosure and complete withholding. This granular negotiability is what makes high-quality, domain-specific data accessible under conditions both parties can accept. This works because data space governance operates at the framework level: it defines how negotiations are conducted, not what is agreed within each one. The Dataspace Protocol (DSP) provides the technical foundation for these negotiations. Domain-specific consortia define the common rules within which they take place. Shared open standards for identity, access rights, and policies ensure these rules hold across organizations. For AI development, this opens access to curated, domain-specific datasets with verifiable provenance. Data that neither web scraping nor generic platform terms of service can reach.

The relationship also runs the other way. Building a data space is largely a matter of integrating its components into the existing backends of participating organizations, and AI substantially reduces the cost of that work by generating metadata, aligning schemas, and monitoring policy compliance across participants. Section 4 develops this direction in full.

2.1 Mutual Enablement: Data Spaces for AI, AI for Data Spaces

The FAIR principles[1] (Findability, Accessibility, Interoperability, and Reusability) provide a practical structure covering the full journey from locating data to using it under agreed terms (Wilkinson et al., 2016). Table 1 summarizes the key interactions.

FAIR-Step	Data spaces for AI	AI for data spaces
Findable	Shared catalogue standards extend the data inventories organizations already maintain internally to the cross-organizational level, allowing automated systems to assess dataset relevance and conditions of use in a single lookup.	AI supports metadata generation for existing datasets being connected to a data space, using empirical data science methods and enabling semantic, vector, and RAG-based search to extend catalogue discoverability.
Accessible	Standardized negotiation and identity protocols extend existing access controls across	The Model Context Protocol (MCP) provides a standardized interface through which AI systems can

[1] Wilkinson, M. D., et al., The FAIR Guiding Principles for scientific data management and stewardship, Scientific Data 3, Article 160018, 2016. https://doi.org/10.1038/sdata.2016.18

FAIR-Step	Data spaces for AI	AI for data spaces
	organizational boundaries, allowing AI systems to obtain governed data access without human intermediaries and serving as a shared compliance control layer over existing data infrastructure.	connect to heterogeneous data services, with MCP servers exposable directly in data catalogues, reducing integration overhead across diverse technical backends.
Interoperable	Open domain standards extend the vocabulary and schema work organizations already manage internally to the cross-organizational level, reducing both the manual overhead of bilateral data sharing negotiations and the computational cost of format translation for AI systems.	AI assists in aligning vocabularies and data models across the heterogeneous backend systems of participants, supporting the development of shared ontologies and bridging terminology differences without requiring full upfront schema standardization.
Reusable	Machine-readable usage policies formalize existing internal governance rules at the cross-organizational level, allowing each rights holder to declare permitted uses per context and enabling enforceable data sharing that preserves data sovereignty.	AI assists in translating legal terms and licenses into machine-readable policies and supports monitoring of policy compliance across participants and processing pipelines. AI can also support the generation of high-fidelity synthetic datasets that preserve relevant statistical and semantic properties of real data, enabling model training, testing, and product development within data spaces without requiring exposure of the underlying sensitive or proprietary data.

Table 1: Data spaces and AI through the FAIR lens

Three cross-cutting points run beneath the row-by-row interactions in Table 1. First, the cost reduction operates at two levels. At the organizational level, shared governance standards remove the need to renegotiate how data is described and exchanged for every new bilateral relationship. At the system level, harmonized schemas cut the computational overhead of format translation, so fewer resources, including the tokens an AI system consumes, are spent reconciling competing standards and more remain for the analytical work itself.

Second, machine-readable usage policies do more than formalize existing restrictions. They enable sharing that was previously impossible under generic platform terms. The web was not built for this: mechanisms such as robots.txt were never designed to govern how data is used once accessed, and large-scale collection of web content for AI training has proceeded largely without rights holder consent, generating significant legal uncertainty (Buick, 2025).

Standards such as the Open Digital Rights Language (ODRL)[2] enable rights holders to specify the permissions, prohibitions, and obligations associated with a dataset in a given context. Enforcement can take several forms. Audit logging provides traceability at the transaction level. Confidential computing environments operated by trusted third parties protect data during processing. In compute-to-data architectures the algorithm is sent to the data holder rather than the data being moved, so data sovereignty is preserved. The European Health Data Space Regulation (EU 2025/327)[3] mandates this model for secondary use of health data, and data spaces such as genome.de and sphin-X are implementing it in practice.

Third, on the AI for data spaces side, one caveat shapes where AI can responsibly sit. LLM-based approaches often lack the reproducibility that production systems require. The canonical role of AI is therefore to assist in developing the fixed ontologies and policies that themselves provide reproducibility, rather than to make runtime decisions that must remain deterministic. Early pilots show both directions at work: a Data Set Search Engine applies empirical and semantic methods to catalogue discovery, and a data space compliance monitor translates legal terms and licences into machine-readable policy (33).

AI can be integrated into a data space in several ways, depending on the role it plays and the governance requirements that apply:

Integration Mode	How it works	Where governance sits
Model as data asset	A model is registered in the catalogue as a shareable asset and negotiated like any dataset, with entries linking it to its training-data provenance and conformity certificates. In compute-to-data scenarios, the model is sent to the data holder rather than the data being moved, often inside a confidential computing environment.	Catalogue entry, usage policy, provenance record and certificates.
AI service via API	An inference service is declared in the catalogue as an endpoint, with its location, interface description, and an application profile defining access conditions (e.g., DCAT application profiles, Swagger/OpenAPI, MCP). A consumer negotiates access through the data space protocol and then invokes the service directly.	Application profile and negotiated access through the data space protocol.
AI embedded in connector extensions	AI capabilities are embedded directly in connector extensions, for instance the Eclipse Dataspace Connector (EDC), enabling automated policy negotiation, real-time metadata enrichment, or AI-assisted access	The connector itself, on the control plane.

[2] https://www.w3.org/TR/odrl-model/
[3] https://eur-lex.europa.eu/eli/reg/2025/327/oj/eng

Integration Mode	How it works	Where governance sits
	decisions without a separate service deployment.	
Processing pipeline declared in the catalogue	Datasets, models and services are declared as a structured pipeline entry and invoked as a single reproducible unit, suited to workflows whose steps must be auditable end to end or orchestrated across many stakeholders (e.g. Piveau, MLflow, Airflow and dbt with metadata represented in DCAT).	Pipeline description bound to the control plane, covering compliance and responsibilities.
Federated learning across participants	Multiple participants train a shared model without raw data, leaving their organizational boundaries. Particularly relevant where data cannot leave regulated environments, such as healthcare or finance.	Terms of participation, who may contribute, under what conditions and how the model is shared, set by the data space.
Governed data for external inference	An external model is grounded at inference time on data retrieved from the data space, for example through retrieval-augmented generation (RAG) over governed catalogue contents and linked data. The model gains context without the underlying records leaving the provider's boundary.	Provider usage conditions, with the retrieval recorded.
AI agent as participant	An AI agent becomes a first-class participant in the data space by interacting through a connector, acting on behalf of its operating organization to autonomously discover, negotiate, and consume data at machine speed. Developed in the next box.	The agent's delegated identity and associated usage policies are bound to an accountable legal entity.

Table 2: Integration modes of AI in a data space

The integration of AI agents into data spaces raises questions beyond technical connectivity4, concerning how agents connect, how they are identified, and how their behavior is governed. In practice, an agent does what a human participant would do, but at machine speed: it discovers data, negotiates access, and consumes assets. Two connection patterns recur. In data space first, the agent enters through a managed third party that handles compliance on its behalf and is not itself a participant. In agent first, the agent acts as a full participant and negotiates directly through the data space protocols, which suit dynamic scenarios with changing data sources. Both use the same underlying protocols; what differs is where access

[4] Connectivity is established through a connector that implements the Dataspace Protocol (DSP), which is currently undergoing international standardization at ISO/IEC as ISO/IEC DIS 26450, Information technology, Eclipse Dataspace Protocol (DSP).

and control are orchestrated. Persistent identifiers give agents stable references to datasets across data spaces, so distributed infrastructures can be navigated as one. The identity and credential mechanisms that make this accountable are detailed in Section 3.

2.2 How AI and data space concepts match

This section maps the conceptual overlap, so readers from either community can recognize the problems they already work on. It establishes the common vocabulary used throughout the paper.

Concept	AI context	Data spaces context	How data spaces help
Digital identity	• IDs of agents • IDs of actors behind the agent • IDs of the datasets used to train the model	• IDs of parties accessing or participating in a data space • IDs of data space offerings • IDs of components like connectors • Verifiable Credentials, the Decentralised Claim Protocol, trust anchors	Both bind an actor to a verifiable identity; data spaces additionally require every agent identity to resolve to an accountable participant, whereas AI identifiers are often ephemeral or self-asserted.
Policy	• intended use/purpose specification, • system prompts and tool-use constraints as operative “policy”, • guardrails	• Machine-readable and technically enforceable • Human-readable and contractually enforceable • Aspirational shared principles of fairness, etc.	The same three registers (enforceable, contractual, aspirational) appear on both sides; they differ in locus of enforcement, with AI tending to embed policy in the model (prompts, guardrails) and data spaces externalising it into independently enforceable, machine-readable artefacts (e.g. ODRL).
Provenance & traceability	LLMOps observability: tracing, logging and real-time monitoring of agent behaviour	Data spaces distinguish backward-looking provenance and data lineage from forward-looking observation of interactions,	Both sides need to know where data came from and went, but training folds data into model weights and largely severs source-to-output

Concept	AI context	Data spaces context	How data spaces help
		supporting accountability and auditability.	lineage, whereas data spaces preserve discrete, transaction-level traceability. Data spaces can therefore supply the provenance AI cannot easily retain on its own.
Data	• Training data • Fine-tuning data • Input data • Output data	• Data products • Datasets • Metadata • Data space offering	AI classifies data by pipeline role (training, fine-tuning, input, output); data spaces by governance and economic status (dataset, data product, offering). The axes cross-cut, so a data product becomes training or input data once consumed.
Platform	AI systems are agnostic over the underlying architecture of the data sources and can work with data platforms and data spaces.	Data spaces are not unitary data platforms or databases and therefore place certain requirements on the AI making use of them.	AI is largely architecture-agnostic and consumes from a platform or a data space alike. Because data spaces are deliberately not a single platform, they let AI reach many providers under one governed set of rules instead of being tied to one platform.
Trust	• "Trustworthy AI" is a normative goal • trust in model outputs is largely probabilistic/reputational/benchmark-based; • trust in an agent rests on its operator, guardrails and track record	Trust is established institutionally and verifiably via attestations, Verifiable Credentials and trust anchors, so previously unknown parties can interact without bilateral pre-agreement.	Both communities centre trust. Data spaces make it verifiable and delegated to an accountable entity, where AI trust is often only statistical or behavioural, so the trust an agent relies on can be checked.

Concept	AI context	Data spaces context	How data spaces help
Contracts and negotiation	Agentic automated negotiation of terms.	Contract negotiation via the Dataspace Protocol; terms are negotiated per dataset and intended use within a framework that governs how, not what.	Data spaces give agents a structured, auditable negotiation surface; AI gives the negotiation autonomy and speed.
Sovereignty	Tendency toward aggregation and centralisation (training corpora, model weights); "control" usually means access controls on a central store.	Data providers control their assets through identity, access and usage policies.	AI's default gravity is centralising, while data spaces are built to preserve distributed control, so participants keep sovereignty over data and models that AI would otherwise pull into one place.
Accountability	Open question of who answers for an autonomous agent's actions.	Actions are traceable to an accountable legal entity through identity and provenance; the agent inherits its operating participant's rights and obligations.	Both need a line from action to responsible party; data spaces supply the mechanism, credentials plus lineage, that AI governance increasingly requires.
Economic value	• Pricing of data and inference • compensating data providers • data valuation for training	Registering and discovering offerings, marketplace functionality and monetisation of data sharing with usage accounting/billing.	Data spaces provide the accounting and settlement layer that makes provider compensation enforceable rather than aspirational.
Boundary component	An MCP (Model Context Protocol) server is the boundary component a system implements once to expose tools, data and services to AI clients; the client discovers and invokes those	A connector is the software a participant deploys to join a data space. Implemented once, it speaks the Dataspace Protocol to every other participant and carries identity, policy	Each is the once-implemented boundary component that lets its side reach a wider ecosystem through a shared protocol instead of point-to-point integration. The

Concept	AI context	Data spaces context	How data spaces help
	capabilities through a single standardised interface rather than bespoke per-service wiring.	enforcement and contract negotiation at the boundary.	difference is governance: a connector authenticates the participant and enforces access policies, usage policies and contracts, whereas MCP standardises discovery and invocation only. The two compose rather than compete, since an MCP server can sit within or behind a connector and gain the governance layer it otherwise lacks.

Table 3: How AI and data space concepts correspond

2.3 Where data spaces complement and improve the AI stack

A general AI stack assumes the data is already at hand and that access was settled somewhere else. It optimizes for model quality and latency, not for who owns the data, on what terms it may be used or how a result can be reproduced and audited later. A data space supplies exactly those missing properties. It gives every dataset a known and accountable provider, a machine-readable description, an explicit usage policy, a verifiable identity for each participant and a provenance and clearing record of what was exchanged. These are not features bolted onto an AI stack after the fact. They are the properties, a data space is built around, and they are what make AI workloads accountable across organizational boundaries.

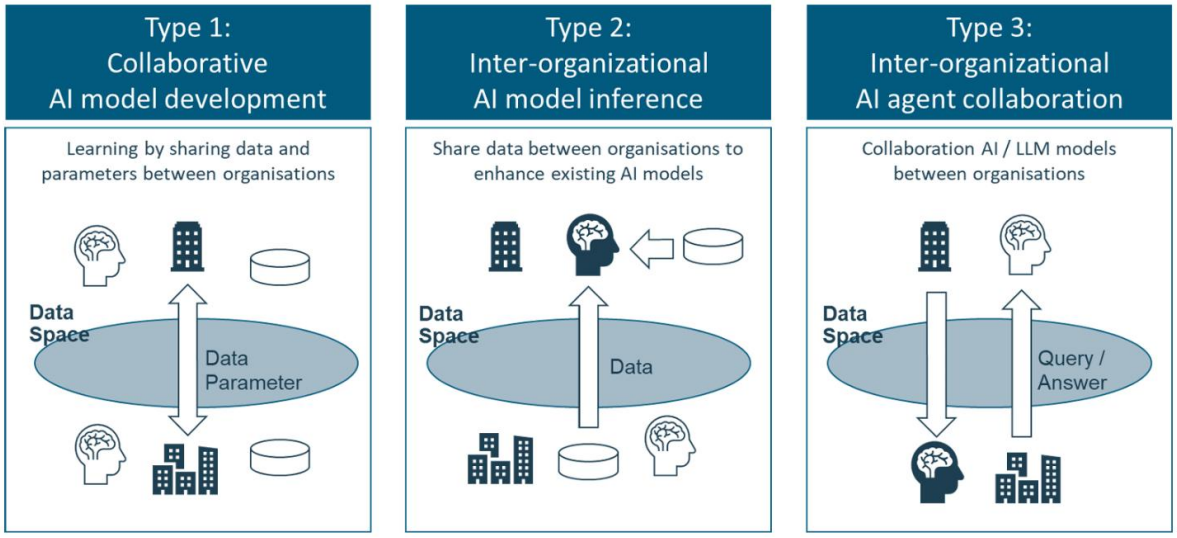

Figure 1: Types of collaborative AI technologies

Fujitsu Research and Fraunhofer ISST[5] describe three patterns by which an AI workload draws on a data space, shown in the figure above. Each pattern places a different demand on the data space, and the data space answers each one with the same governed primitives.

In the first pattern, **Collaborative AI Model Development**, organizations train a shared model by exchanging data or model parameters rather than keeping each model in isolation. Here the data space contributes the usage policies that decide which data may enter training and for what purpose, together with the provenance record that documents which contributions shaped the resulting model. This is the basis for sovereign federated learning, where the training data never leaves the contributor, but its agreed contribution still counts.

In the second pattern, **Inter-organizational AI Model Inference**, one organization keeps its own model and enriches it at inference time with data held by others, for example through federated retrieval augmented generation. Here the data space contributes catalogue discovery, access control and the usage conditions that bound how a retrieved answer may be used, so that context can be pulled across organizations without exposing the underlying records.

In the third pattern, **Inter-organizational AI Agent Collaboration**, each organization already runs its own model or agent, and the parties accomplish a task by having those agents query and negotiate with one another. Here the data space contributes participant identity, contract negotiation and the audit record, so that an autonomous exchange between agents of different organizations stays bound to accountable parties and remains observable.

The three patterns are not exclusive and a single deployment can combine them. They recur in the pilots in 33, tagged with the matching pattern where one applies. The reverse direction, in which AI serves the operation of the data space itself rather than consuming it, is developed in Section 4 and illustrated in 5.1.

As AI systems become autonomous participants, sovereignty extends beyond protecting data to governing AI-mediated decisions. Data sovereignty ensures control over information

5 Source: Fujitsu Research in cooperation with Fraunhofer ISST, 'Decentralized and Collaborative AI for Data Spaces,' 2025 (reused with permission)

assets, while decision sovereignty ensures that organizations retain visibility and control over how autonomous decisions are generated. Data space mechanisms such as verifiable credentials, policy enforcement, provenance records and auditable interaction histories provide the foundations for this capability. Together they make it possible to determine under which policies an AI acted, which trusted datasets or models contributed to a decision and how delegated interactions evolved over time. This level of transparency becomes increasingly important as organizations exchange not only data but AI-enabled capabilities across organizational boundaries.

3 Today's AI Challenges and how data spaces solve them

3.1 Trust stated as the challenge

Trust is the precondition for everything else in this paper. An organization will let an external agent act on its data only if it can verify, before any action, that the agent, the data and the services involved are what they claim to be. In a data space this verification is automated. Verifiable Credentials are the operative mechanism at runtime: fitness attestations, accreditation proofs and conformity certificates travel with the asset they describe, so a participant can check them at the point of use without retrieving documents by hand and without gaps in the audit chain.

These credentials are organized into a trust hierarchy, which is what lets each check be verified locally without re-involving whoever issued it. At the base sits institutional accreditation, for example by a national accreditation body such as the German National Accreditation Body, Deutsche Akkreditierungsstelle (DAkkS). Accreditation qualifies bodies to issue fitness attestations for tools and datasets, at an agnostic level through testing tools such as Quality-X and QI-Digital and at a domain level through bodies such as sphin-X for health. Those attestations are expressed as Digital Certificates of Conformity (D-CoC) carried as machine-readable Verifiable Credentials. Running through all layers, SMART-standards (IDIS[6]) provide a shared normative ground truth: requirements are delivered as persistently identified, machine-readable units that flow unchanged into attestations and compliance artefacts, which prevents drift and hallucination at every layer. On this basis a domain body such as sphin-X can act as a Qualified Body for health, issuing attestations for health-AI testing tools, curating reference datasets and producing certificates that regulators and procurement bodies can consume directly through the data space.

This hierarchy makes two distinct things verifiable. The first is the AI itself. The quality and conformity of a model become transparent through data space credentials, and a bill of data, a bill of software and a bill of licences can document what went into its training, making the trustworthiness of the sources auditable. The second is data access. Participants gain full transparency over what data may be used for what purpose, which removes any excuse for training on data without permission and directly supports IP protection. Because the trust framework and the Dataspace Protocol make these terms explicit and enforceable, they also make data available that generic platform terms would have kept closed.

AI services are themselves offered as data space assets, with models published alongside the context information needed to assess them. Here the standardization is still maturing. There is not yet an agreed way to declare application profiles or to publish a model and its

[6] DKE, SMART standards (IDIS). https://www.dke.de/idis-en

automated tests as governed assets, and closing that gap is one of the open items for the standards community.

3.2 Data quality and availability

One of the most persistent challenges for AI is not a lack of algorithms but a lack of reliable, well-described and legally usable data. AI systems depend on data that is accurate, representative, timely and sufficiently complete for the purpose, yet organizations face fragmented silos, inconsistent formats, missing metadata, uncertain provenance and unclear rights to reuse. These weaknesses feed directly into AI behaviour, producing biased outputs, hallucinations, low robustness and decisions that cannot be explained or audited.

The problem is sharper for cross-organizational AI, because many high-value use cases need data that no single organization holds, for example predictive maintenance across supply chains, energy flexibility services or healthcare research. The relevant data sits with competitors, public authorities, suppliers and research institutions, and even where it exists it is often withheld for reasons of confidentiality, commercial sensitivity or privacy.

A data space addresses this by creating a governed environment for trusted sharing without centralizing the data. Participants make datasets discoverable through standardized metadata and catalogues while keeping the data itself under their own control, so potential users can find datasets, understand their characteristics and request access under defined conditions, and providers retain sovereignty over who may use the data, for what purpose and on what terms.

It also raises the quality of what is shared. Each dataset carries its provenance, collection method, update frequency, quality indicators, semantics, usage constraints and licensing, which is the context a developer needs to judge whether it is fit for a given purpose. Treated this way a dataset becomes what is increasingly called a data product: not raw signals but information refined for reuse, carrying the content, quality, context and machine-readable format that a model or application can consume directly. What counts for AI is then fitness rather than volume, and the data space is where fragmented signals from systems, machines and supply chains are turned into governed, reusable products. Shared vocabularies and semantic models let those products be combined across sources without losing meaning.

Finally, usage-control policies, identity, certification and logging let providers share selectively rather than exposing data without control, which unlocks collaboration that would otherwise be impossible: data accessed only under strict purpose limitation, processed only in a secure environment, or combined with privacy-preserving techniques such as federated learning or synthetic data.

3.3 Access, usage conditions and selection of data

Authorizing an AI agent in a data space means bridging static human identity and dynamic machine execution. Traditional Identity and Access Management was built for human users and fixed service accounts, and does not capture an autonomous agent whose model, tools and purpose can change between sessions.

The data space answers this with a Blended Digital Identity: a Participant ID that cryptographically combines the autonomous agent with its legally liable organization. This binding is what sets the criteria for selecting a permissible agent for a given set of data, taking

account of data type, location and usage policies, and it is the primary defence against unauthorized or malicious AI entering the data space.

Verifiable trust requires distinguishing three identity layers, each with its own lifecycle and accountability owner:

Legal Participant Identity	Agent Identity	Tool and Service Identity
The static, legally anchored identity of the enterprise or individual deploying the agent, verified for example via eIDAS2 or global corporate registries. The lifecycle is permanent, and accountability rests with the corporate entity.	The transient, software-defined execution layer. Its lifecycle is dynamic, spun up for a specific task and spun down on completion. Accountability rests with the human supervisor or internal system that authorized the deployment.	The specific capabilities, tools and endpoints the agent is authorized to use. Its lifecycle is tied to version control and security updates, with accountability resting on the software developer or AI service provider.

Table 4: Three Layers of Agentic Identity

These three layers are bound together in a Verifiable Credential called the Delegated Agent Participant ID, the credential that links an executing agent back to its legally liable organization. Before any data transaction begins, the data space connector verifies six elements within it: the organization, as the cryptographically verified Decentralized Identifier of the legal entity; the agent instance, a transient runtime identifier for the specific session; the model powering the agent's reasoning; the tool registry profile listing the external APIs and capabilities the agent may use; the purpose scope, a machine-readable declaration of intent that must align with the provider's ODRL usage policies; and the assurance level, the agent's trust tier backed by testing certificates or accreditations.

Autonomous collaboration adds one further requirement. When an authorized agent needs specialized data outside its immediate context, it may task a secondary agent across organizational boundaries, and the data space governs this through cascading delegation. The secondary agent inherits the boundaries of the primary agent's Delegated Agent Participant ID, in particular its purpose scope and assurance level, so usage policies, sandboxing constraints and legal liability flow down the chain. No sub-agent can reach data or take actions the primary organization was not authorized to handle.

3.4 Compensation of data providers

In traditional data space architectures, the compensation of data providers relies on human-mediated economic capabilities, such as flat-fee subscriptions, bulk licensing agreements, or manual invoicing. While sufficient for static data sharing, these traditional mechanisms create an absolute bottleneck for Agentic AI.

As AI models transition into autonomous actors, they consume data at machine speed and in highly specific micro-increments. To support this velocity, the data space must evolve

beyond simple data transfer and develop the foundational capability for machine-to-machine (M2M) economic settlement, orchestrated by the ecosystem's specialized agents.

Dynamic Pricing Capabilities

To incentivize organizations to share high-value, proprietary data, the ecosystem requires the capability to price assets programmatically rather than statically. This capability begins with the cataloguing and curation agent, which must be able to embed flexible commercial rules and usage-based pricing conditions directly into the data offering's metadata.

On the consumer side, the discovery (scout) agent and data-gap analysis agent require the capability to autonomously evaluate these dynamic costs against the organization's operational ROI before initiating a transaction.

Finally, the negotiation agents (representing both provider and consumer) must possess the capability to autonomously propose, counter-propose, and resolve pricing structures in real-time, utilizing supply-and-demand insights surfaced by the matchmaking (broker) agent.

Machine-to-Machine (M2M) Settlement Capabilities

To facilitate high-frequency micro-transactions, data spaces must provide the architectural capability to transfer value instantaneously across organizational boundaries, eliminating the administrative overhead of traditional clearing.

Once the conformity and compliance agent exercises its capability to verify that a transaction meets all regulatory and policy boundaries, the purchasing negotiation agent must have the capability to autonomously authorize and trigger compensation from a pre-approved corporate budget. The data space infrastructure must be capable of processing this settlement concurrently with the data request.

Atomic Execution and Metering Capabilities

The most critical economic capability required for an Agentic Mesh is the "atomic transaction", the ability to ensure that the authorization of data access and the compensation of the data provider occur as a single, indivisible event.

To support complex, usage-based economic models (such as paying per query or per reasoning task), the provenance and observability agent must possess the capability to act as an automated, ecosystem-wide meter. It requires the capability to securely track, measure, and log the exact volume of data or computational resource consumed during an agentic workflow, ensuring that data owners possess indisputable proof of usage to guarantee precise, automated compensation.

3.5 Observability and traceability

A key challenge for trustworthy AI is the difficulty of observing and tracing data across distributed ecosystems. AI systems often rely on data collected, transformed, enriched, and reused by multiple actors. Once this data is used for training, validation, retrieval-augmented generation, or automated decision-making, it can become difficult to reconstruct where the data came from, how it was modified, who accessed it, and whether it was used according to agreed conditions.

This is where data spaces provide a strong response. ISO/IEC 20151-1, Cloud computing and distributed platforms — Dataspaces — Part 1: Concepts and characteristics[7], defines foundational concepts and essential characteristics of data spaces, emphasizing trusted data sharing between participants under agreed governance, policies, semantic models, protocols, processes, and enabling services. In this context, observability and traceability are not optional add-ons, but core operational capabilities of a trustworthy data-sharing environment.

For AI, traceability enables organizations to link model behaviour and AI outputs back to the data sources, access conditions, transformations, and usage policies that shaped them. Observability provides visibility over data flows, access events, policy enforcement, quality indicators, and anomalies during operation. Together, they support auditability, accountability, incident investigation, regulatory compliance, and continuous monitoring of AI systems.

Data spaces therefore help turn AI governance from a static documentation exercise into an operational capability. By embedding identity, metadata, provenance, logging, usage policies, and governance mechanisms into cross-organizational data sharing, they allow participants to demonstrate not only that data was shared, but also how, by whom, under which conditions, and for what purpose.

3.6 Governance, automation and the data space governance model

Every data space needs a governance authority: the body that maintains its rulebook, admits and removes participants and oversees compliance. This authority takes one of two forms. It can be a dedicated legal entity, such as an association, foundation or company, which holds the shared infrastructure and carries liability in its own name. Or it can be a lighter contractual arrangement, in which the participants sign one shared agreement and govern themselves through a steering committee without creating a separate organization. Whichever form it takes, the authority's own rules operate inside the wider regulatory environment and the legislation that defines what is permitted.

This governance now has to account for AI agents, both those acting inside the data space and those of outside organizations seeking to discover and negotiate data from it. Each agent has to be recognized as an actor in its own right, with a defined role and bound by the identity and accountability mechanisms set out in 3.3. The Dataspace Protocol supports this at the protocol level by separating two planes, shown in Figure 2.

[7] ISO/IEC 20151, Cloud computing and distributed platforms — Dataspaces — Part 1: Concepts and characteristics. https://www.iso.org/standard/86589.html

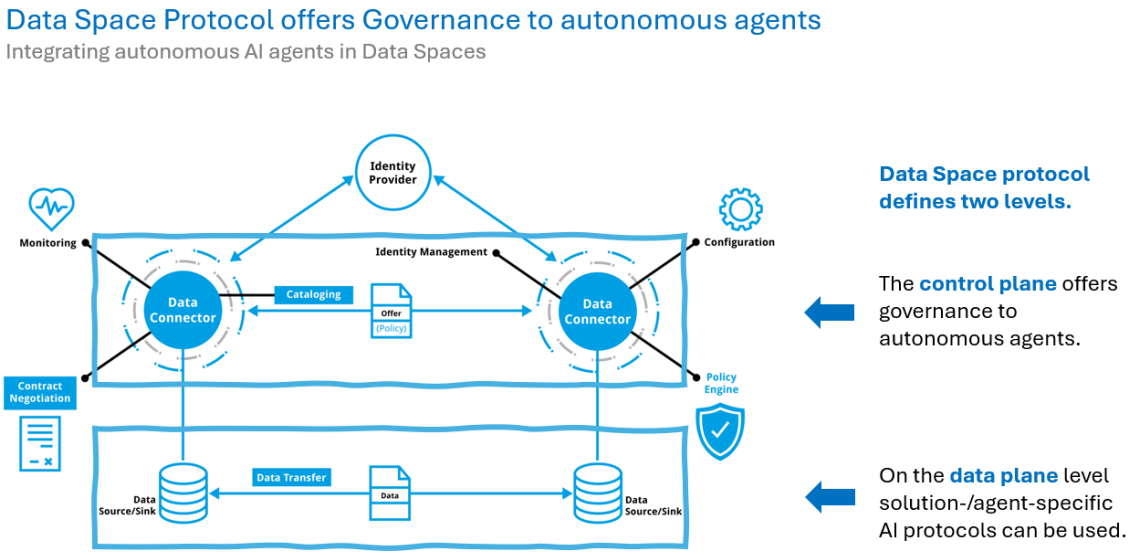

Figure 2: Dataspace Protocol defines two levels. The control plane provides governance to autonomous agents; on the data plane, agent-specific AI protocols such as MCP and A2A can be used. Source: Achatz (2025), IDSA

The control plane carries governance: identity, access negotiation and the policy decisions that keep every agent action attributable and authorized. The data plane is where the exchange itself happens and there, agent-specific protocols such as MCP and A2A can operate within the boundaries the control plane has set. Keeping the two apart is what lets agent tooling evolve quickly on the data plane without loosening the governance that holds on the control plane. In the longer term this is the basis for machine-readable governance, where agents join, negotiate and consume data through these protocols without an organization joining each data space by hand.

3.7 Agent discovery and integration

In a data space, agents are discovered through the same catalogue that lists data and services. An agent can be registered as a catalogue asset, which makes its capabilities visible and negotiable: its entry declares what the agent does and carries the ODRL policies that govern it, including when it may act autonomously, when payments to other providers are authorized and when human approval is required. A participant or another agent can therefore locate a suitable agent, review its terms and negotiate access without any prior bilateral arrangement, so discovery and governance happen in a single step.

Integrating with existing applications rests on the same foundation. Many applications already rely on data spaces for trusted exchange, and a data space gives them a standardized way to connect so their data can be used and, in turn, reached by AI agents. This builds on the emerging data space standards: ISO/IEC 20151 defines the underlying concepts, and the Dataspace Protocol, undergoing international standardization as ISO/IEC DIS 26450, defines how participants connect and negotiate. In practice, integration means giving the agent a connector as its governed entry point to the data space and the applications behind it. The connector handles the data side. Running the agents themselves needs a second layer for model access, tool integration, orchestration, policy control and observability, sometimes called an agentic hub, which reuses the data space's identity, access and usage mechanisms rather than building its own, so an agent's actions stay as accountable as the data it consumes.

3.8 Ethical and societal expectations

AI deployed across organizational boundaries raises questions about fairness, accountability plus the wider impact on society. This section frames the general risks of using AI in cross-organizational settings. It explains the trust gap that data spaces are designed to close. Efforts across politics, business and civil society to make ethics applicable to AI are increasingly grouped under the heading of digital ethics: the branch of ethics concerned with moral problems relating to data, algorithms and the surrounding practices and infrastructures (Floridi et al., 2018)[8]. The field divides into the ethics of data, the ethics of algorithms and the ethics of infrastructure. A data space does not settle the first two for its participants, but it sits squarely in the third. The ethics of infrastructure concerns the responsibilities of those who design and operate data processes, and a data space is the infrastructure through which those responsibilities can be made explicit, assigned to an accountable party and checked.

These risks are not new to AI but deploying it across organisational boundaries sharpens every one of them. The EU High-Level Expert Group sets out seven requirements for trustworthy AI: human agency and oversight, technical robustness and safety, privacy and data governance, transparency, diversity and non-discrimination, societal and environmental well-being, and accountability.[9] Each names a distinct way a system can fail the people it affects, by acting without meaningful human control, behaving unpredictably, mishandling personal data, concealing how it reached a result, treating groups unequally, imposing costs on society or the environment, or leaving no one answerable when something goes wrong. Inside a single organisation, one party can in principle check all seven against its own data and systems. Across organisational boundaries this breaks down. An AI system may be trained or run on data whose origin, quality, and permitted uses are controlled by someone else, and the organisation deploying it cannot vouch for what it cannot see.

This is the trust gap, and it is worth being precise about what kind of gap it is. It is less a shortage of goodwill than a shortage of verifiability: each party can act in good faith and still have no way to prove to the others that it has. Data spaces are designed to close this specific gap. Their founding principles, set out in the IDSA Manifesto,[10] treat trust as something to be built into infrastructure rather than assumed: data spaces are described there as “a mechanism to create trust,” and, more pointedly, participants are asked to “act in good faith, but verify.” The same principles explain how the gap is closed. Because data does not flow through a central platform and no operator sits above the rest, control stays with each provider (“your data, your choice”), while shared standards let independent parties check one another’s claims without submitting to a common master. What this buys for AI is real but bounded. A data space cannot make a model fair or a dataset unbiased; it can make the provenance, quality, permitted uses, and accountability of the data verifiable to everyone who relies on it. That turns several of the trustworthy-AI requirements from promises into properties that can be checked.

The wider stakes are societal, and here the two agendas point in the same direction. The EU Declaration on Digital Rights and Principles asks that digital systems put people and their rights at the centre, support solidarity and inclusion, preserve freedom of choice, foster

[8] Floridi, L. (2018). Soft ethics and the governance of the digital. Philosophy & Technology, 31(1), 1–8. https://doi.org/10.1007/s13347-018-0303-9
[9] European Commission, High-Level Expert Group on AI, Ethics Guidelines for Trustworthy AI, 2019. https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai
[10] International Data Spaces Association, The Data Space Manifesto, Version 1.0, April 2025. https://internationaldataspaces.org/wp-content/uploads/dlm_uploads/The-Data-Space-Manifesto-Version-1.0-April-2025.pdf

participation, keep individuals safe and empowered, and sustain the digital future.[11] An architecture that leaves people and organisations in control of their own data, rather than concentrating it in a few hands, is structurally aligned with these aims. But alignment is a possibility, not a guarantee, the manifesto's own warning that "with great responsibility comes great power" cuts both ways, and the instruction to verify applies to data spaces themselves as much as to anything built on them. The contribution of bringing data spaces and AI together is therefore not that it makes technology ethical, but that it makes societal expectations auditable: it gives the people affected, and those acting on their behalf, a way to see whether those expectations are being met.

Data spaces do not automatically guarantee that shared data is accurate, unbiased, or fit for purpose. However, they provide governance and technical mechanisms that make claims about data more transparent, auditable, and verifiable. This is especially relevant for AI use, where consumers, whether human users or AI agents, need to understand not only the content of a dataset, but also its origin, quality, limitations, permitted uses, and whether it represents original, transformed, aggregated, or synthetic data. Where synthetic data is shared in a data space, its synthetic nature can be explicitly declared through metadata, provenance information, usage policies, and quality indicators, enabling consumers to assess its suitability without assuming equivalence with the original data.

3.9 Regulatory landscape worldwide

AI regulation varies widely across jurisdictions, from binding horizontal law to promotion-oriented soft law. The common thread for this paper is that whatever the regulatory style, data spaces supply the operational mechanisms (identity, usage policies, provenance and audit) that turn governance expectations into verifiable practice. The table below summarizes the main approaches and where data spaces help.

Jurisdiction	Approach to AI governance	Where data spaces help
European Union	Binding, risk-based horizontal regulation: the EU AI Act (Regulation 2024/1689)[12], layered over the GDPR (Regulation 2016/679) and sector-specific rules.	Identity, machine-readable usage policies, provenance and audit logs make AI Act and GDPR obligations such as data governance, documentation and traceability demonstrable in practice rather than asserted.
Japan	Promotion-oriented soft law: the AI Promotion Act (Act No. 53 of 2025) sets basic principles rather than penalties, with best-effort duties and non-binding guidance such as the AI Guidelines for Business, alongside AISI evaluation methods.	Soft law places the burden of trust, provenance and usage control on infrastructure rather than on penalty-backed regulation, the role data spaces are built for. Japan

[11] European Commission, European Declaration on Digital Rights and Principles. https://digital-strategy.ec.europa.eu/en/policies/digital-principles
[12] European Union, Regulation (EU) 2024/1689 laying down harmonised rules on artificial intelligence (Artificial Intelligence Act), 2024. https://eur-lex.europa.eu/eli/reg/2024/1689/oj

Jurisdiction	Approach to AI governance	Where data spaces help
		couples this to a national data space strategy.
China	Promotion-oriented Action with scenario-based regulations: the State Council published "AI+ Action" in 2025, as a top-down and pro-innovation action plan on AI development integrating 6 fields, while CAC as a regulator issued rules and guidelines on developed scenarios such as AIGC, facial recognition, AI agent applications.	Data spaces embed the compliance requirements of AI governance through data usage control and audit traceability, serving as the data transferring infrastructure that enable the systems to be Design-as-Compliance.
Republic of Korea	The AI Basic Act, promulgated January 2025 and effective January 2026, both promotes the AI industry and sets a foundation for safe and trustworthy use, with a grace period of at least one year and some transparency penalties deferred. [13][14][15]	The same provenance and accountability mechanisms support the Act's trustworthy-use foundation.
Brazil	Bill PL 2338/2023 is the central AI legislation, processed through 2025, while the LGPD remains the core personal-data framework. [16][17]	Data spaces supply the access control and usage accounting that AI use of personal data under the LGPD requires.
India	A principle-based, techno-legal approach (India AI Governance Guidelines) covering data management, transparency, risk classification and safety testing, with the DPDP Rules 2025 governing digital personal data. [18][19][20]	Provenance, usage policies and auditability operationalise these principles across organizational boundaries.

[13] Ministry of Science and ICT (MSIT), Republic of Korea, AI Basic Act: passage and promulgation. https://www.msit.go.kr/eng/bbs/view.do?bbsSeqNo=42&mId=4&mPid=2&nttSeqNo=1071&sCode=eng
[14] Ministry of Science and ICT (MSIT), Republic of Korea, AI Basic Act: industry promotion and a foundation for safe and trustworthy AI. https://english.msit.go.kr/eng/bbs/view.do?bbsSeqNo=42&mId=4&mPid=2&nttSeqNo=1238&sCode=eng
[15] Ministry of Science and ICT (MSIT), Republic of Korea, AI Basic Act: grace period and deferred transparency penalties. https://www.msit.go.kr/eng/bbs/view.do?bbsSeqNo=42&mId=4&mPid=2&nttSeqNo=1215&sCode=eng
[16] Federal Senate of Brazil, Bill PL 2338/2023 on the use of artificial intelligence. https://www25.senado.leg.br/web/atividade/materias/-/materia/157233
[17] National Data Protection Authority (ANPD), Brazil, Lei Geral de Proteção de Dados Pessoais (LGPD), Lei nº 13.709. https://www.gov.br/anpd/pt-br/centrais-de-conteudo/outros-documentos-e-publicacoes-institucionais/lgpd-en-lei-no-13-709-capa.pdf
[18] Press Information Bureau, Government of India, India AI Governance Guidelines, February 2026. https://static.pib.gov.in/WriteReadData/specificdocs/documents/2026/feb/doc2026215790801.pdf
[19] IndiaAI, Safe & Trusted AI. https://indiaai.gov.in/hub/safe-trusted-ai
[20] Press Information Bureau, Government of India, Digital Personal Data Protection (DPDP) Rules 2025, operationalising the Digital Personal Data Protection Act 2023, November 2025. https://static.pib.gov.in/WriteReadData/specificdocs/documents/2025/nov/doc20251117695301.pdf

Jurisdiction	Approach to AI governance	Where data spaces help
United States, Australia, United Kingdom	Less centralized than the EU: fragmented US federal guidance, voluntary standards, sectoral enforcement and state law; Australia moving from voluntary guidance toward mandatory guardrails for high-risk AI; a pro-innovation, regulator-led UK model.	Across all three, data spaces provide the operational mechanisms (provenance, access control, usage policies, auditability and accountability) that turn high-level expectations into verifiable practice.

Table 5: Approaches to AI governance across jurisdictions, and where data spaces help

Japan is the clearest case of data-space strategy advancing in step with AI policy. Cross-sectoral sharing is coordinated through the Data Society Alliance under the DATA-EX initiative and METI's Ouranos Ecosystem, and in October 2025 IPA, the DSA, the Robot Revolution & Industrial IoT Initiative and the University of Tokyo agreed to promote Open Data Spaces as a neutral technical concept, anchored in an Open Data Spaces reference architecture and protocol aimed at international interoperability with International Data Spaces. The xIPF Consortium (33.2) is the most recent, explicitly AI-native expression of this trajectory.

4 How AI adds value to data spaces

This is the "AI for Data Spaces" arm of dual framing introduced in Section 2.1. Where Section 2 establishes shared vocabulary and Section 3 covers the challenges, Section 4 sets out the concrete value AI brings to operating a data space.

A data space is, before it is anything else, a coordination layer over metadata: catalogues of offerings, vocabularies and ontologies, machine-readable usage policies, participant identities, provenance records and contracts. The data itself is exchanged directly between participants rather than being routed through a central platform. This makes the data space an unusually rich substrate for AI, because the value AI adds comes less from operating on payload data than from reasoning over the dense, structured, governed knowledge that surrounds it. Section 4 sets out that value: the pressures agentic AI places on data space operation (Section 4.1), the roles agents can occupy as super-automated participants (Section 4.2), the capabilities they deliver (Section 4.3) and the trust separation that keeps all of this accountable (Section 4.4).

4.1 Key challenges AI poses for data space operation and governance

Agentic AI is attractive because it is capable. An agent built around a large language model does not just execute fixed instructions. It can interpret a goal, plan a path toward it, call external tools and services, absorb feedback to adjust its actions and collaborate with other agents inside and across organizations. That same autonomy is what challenges a data space. The controls that keep traditional systems in a steady state, access control, data isolation and contractual constraints in closed environments, assume a human pace and a fixed client. Agentic participation breaks both assumptions. The pressures it places on data space operation and governance fall into six types.

Pressure	What it means for data space governance?
Speed	Agents issue far more requests than humans and can act without manual approval, so enforcement must be automated and held at machine speed.
Ambiguity	Natural-language instructions can be underspecified, so an agent may pursue goals it was never authorized to. Purpose and scope must be explicit and policy-checked, not inferred by the agent.
Composability	Agents chain retrieval, reasoning and tool calls, so a permitted first step can lead to an impermissible later one. Policy must be evaluated at every boundary, not only at entry.
Malleability	Behavior shifts with model updates, prompts, tools, context, fine-tuning or adversarial inputs, so the entity being trusted is not fixed. Identity and assurance become harder to pin down.
Epistemic risk	Agents can produce plausible but false outputs, miss constraints or be steered by poisoned context. When output drives access,

Pressure	What it means for data space governance?
	publication or a contractual obligation, this becomes governance risk.
Attack surface	Prompt injections, tool injections, data exfiltration, poisoned knowledge and confused-deputy attacks appear once agents reach external systems, and agents can be impersonated. Enforcement must be deterministic and sit outside the model context.

Table 6: The pressures agentic AI places on data space governance

These six pressures share one root. Autonomous action that can cross organizational boundaries must stay bound to an accountable participant, an explicit purpose, a policy decision and an auditable record, even when the agent is fast, adaptive and only partly predictable. This is neither a new protocol problem nor unexplored ground. IDSA Rulebook already describes [21]how an agent participates on behalf of an organization and how trust and access are established before it acts.

4.2 Agent roles in a data space

Section 2.1 frames AI agents as super-automated humans: whatever a human participant can do through the data space's interfaces, an agent can do through the same interfaces, faster and continuously. The identity model of Section 3.3: the legal participant identity, the agent identity, and the tool/service identity, bound together in a Delegated Agent Participant ID, is what makes this safe: every agent action is attributable to a legally liable organisation. Against that backdrop, several distinct agent roles emerge in data space operation:

Agent	Role
Discovery (scout) agent	Continuously searches the catalogue and knowledge graph for participants, datasets and offerings that match a standing need, surfacing opportunities a human would never have time to find.
Data-gap analysis agent	Compares an organisation's requirements (or a model's data needs) against what it already holds and what the space offers and identifies the gaps worth filling.
Negotiation agent	Conducts the contracting process for data sharing, proposing and counter-proposing usage policies, pricing and terms through the data space's negotiation flow, with a human approving the final agreement (the subject of the pilot below).
Cataloguing and curation agent	Enriches metadata, maps vocabularies and keeps the knowledge graph current.

[21] International Data Spaces Association, AI Agents (AI and Dataspaces: Shaping the Future of Trusted, Decentralized Intelligence), IDSA Rulebook. https://kb.internationaldataspaces.org/external/rulebook/130_AI_Agents/

Agent	Role
Conformity and compliance agent	Checks offerings, policies and the agent's own intended use against verifiable credentials and the regulatory context (cf. Section 3.1, Section 3.9).
Provenance and observability agent	Reconstructs lineage and monitors usage for traceability (Section 3.5).
Matchmaking (broker) agent	Pairs supply and demand across the ecosystem.

Table 7: Agent roles in a data space

These roles compose. A typical workflow chains them: a scout agent finds candidate partners, a gap-analysis agent confirms the data closes a real need, a conformity agent verifies the offering's credentials, and a negotiation agent settles terms, each acting through standard data space interfaces under its delegated identity, mapping cleanly onto both the Dataspace-First and Agent-First interaction patterns of Section 2.1.

4.3 AI capabilities for data space operation

The roles above are delivered by a set of cross-cutting capabilities. The most consequential are:

- **Automated partner and offering discovery.** Semantic search over the knowledge graph lets agents find relevant participants and data products by meaning rather than exact keyword, including non-obvious matches surfaced through graph inference.
- Data-gap analysis. By reasoning over requirements versus available offerings, AI can quantify what is missing for a given purpose: a model to be trained, a product to be assembled and prioritise acquisitions, turning a manual, expert-dependent exercise into a continuous one.
- Automated negotiation of data-sharing contracts. AI agents negotiate the usage policies (e.g. ODRL), pricing and conditions of a data transaction through standardised protocols, lexicons and data formats, converging on mutually acceptable terms while leaving the final commitment to a human. NEC's Automated Negotiation AI illustrates the maturity of the approach: in a 2024 proof-of-concept it reached an automated agreement rate of around 95% and reduced negotiation-to-completion time to roughly 80 seconds.22 Applied inside a data space, this collapses the contracting friction that today limits the number of data-sharing relationships an organisation can sustain.
- Metadata enrichment and semantic interoperability. AI maps schemas, aligns vocabularies and proposes ontology links, lowering the integration cost that interoperability otherwise demands (Section 2.2).

[22] NEC Corporation, Automated Negotiation AI, 2024 proof-of-concept. https://www.nec.com/en/global/solutions/ai/analyze/negotiationai.html

- Policy authoring and enforcement support. Agents help translate human-readable, contractually framed terms into machine-readable, technically enforceable policies, bridging the policy gap identified in Section 2.2 and flag conflicts before they reach enforcement.
- Quality assessment, conformity verification and provenance reconstruction. AI evaluates fitness-for-purpose, checks conformity credentials at point of use (Section 3.1), and reconstructs lineage to support traceability obligations (Section 3.5).
- Natural-language access. Participants express needs in natural language; the agent translates them into catalogue queries, policy checks and negotiation actions, lowering the expertise barrier to participating in a data space at all.

Taken together, these capabilities shift the data space from a system that humans operate transaction by transaction to one that agents operate continuously on participants' behalf, under delegated and accountable identities. That shift is what the pilot in the next section demonstrates in practice.

4.4 Trust formula

Trustworthy agentic AI in data spaces requires a clear separation between probabilistic proposal and deterministic enforcement. The model may propose. The Knowledge Engine checks. The policy engine authorizes. The data space records. The human steward always remains accountable.

This paper proposes the following trust formula as a practical design heuristic:

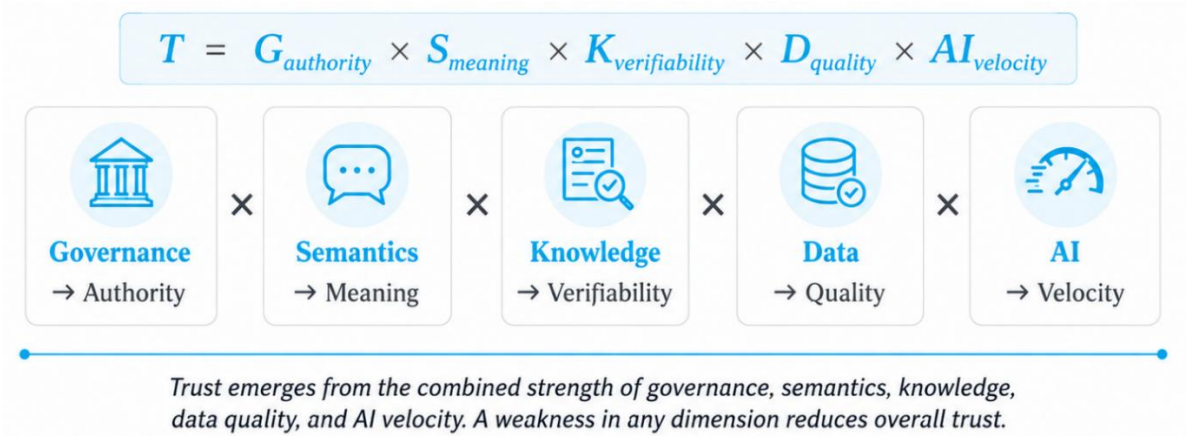

Figure 3: Formula of trust

The formula is multiplicative. If any factor is zero, trust collapses. A fast AI system without policy is not trustworthy. A policy system without semantics cannot decide correctly. A semantic model without data quality cannot produce reliable outputs. A knowledge base without verification can turn errors into executable rules. A data space provides the environment in which these factors can be assembled and enforced.

5 Use Cases

5.1 Cross-sectoral examples

The pilots and proofs of concept in this section show the mutual enablement of Section 2 at work. Some demonstrate data spaces for AI and agentic participation, where agents discover, negotiate and transact on a participant's behalf. Others demonstrate AI for data spaces, where AI serves the operation of the space itself through compliance checking, interoperability support, data curation and improved discovery. Each pilot indicates the interaction pattern it illustrates.

The pilots are grouped by what they demonstrate: agentic participation in and around the xIPF Consortium (5.2 and 5.3), then AI applied to the operation of the data space itself (5.2.3 to 5.2.5).

5.2 The xIPF Consortium

xIPF Consortium is a Japanese industry-academia initiative established as a general incorporated association on April 10, 2026. Under its Articles of Incorporation, its founding members are NTT DATA Group Corporation, SoftBank Corp., East Nippon Expressway Company Limited and Fujitsu Limited. In addition, organizations such as the University of Tokyo (Koshizuka Lab, Interfaculty Initiative in Information Studies), NEC, the Ouranos Ecosystem Promotion Center and the Data Society Alliance (DSA) participate in the consortium.

The consortium's stated aim is the realization of a society in which AI, data spaces and related technologies work in concert, so that AI and data can be used safely and flexibly, enabling large-scale, cross-organizational data use with AI while preserving the security, reliability and sovereignty properties this paper has argued are essential. The envisaged approach is cooperative rather than centralizing: enterprise AI platforms, large language models and organizations' existing data systems are to interoperate across organizational boundaries, with each organization retaining control of its own systems and data. Target application domains include logistics, mobility, energy management and urban planning. In the terms of Section 2, this agenda brings Data Spaces for AI and AI for Data Spaces together: AI is both a primary consumer of shared data and a means by which data sharing is operated.

5.3 The NEC × EverySense Japan prototype: automated negotiation of data-trade terms

Within this context, NEC and EverySense Japan (ESJ) demonstrated a prototype for the automated negotiation of data-trade terms. EverySense Japan operates a data-marketplace platform (EverySense Pro, in operation since 2016) that intermediates the exchange of a wide variety of data between providers and consumers, supplying the marketplace context, the catalog of tradable data and the matching of supply to demand[23]. NEC contributed its Automated Negotiation AI, in which each party is represented by a negotiation agent that searches for mutually acceptable terms using standardized protocols, data formats and lexical definitions.

[23] EverySense, Inc., EverySense Pro data marketplace, in operation since 2016.

In the demonstration, a data trade is negotiated fully automatically between two participants, a data buyer and a data seller, each represented by its own negotiation agent. The buyer discovers a data source. The buyer's agent initiates contact with the seller's agent, and the two agents negotiate the terms of the transaction without human intervention, converging on a mutually acceptable set of data-trade terms across three dimensions:

- Data characteristics: the physical quantities supplied (temperature and pressure readings in the demo).
- Data-sharing conditions: the usage terms governing the shared data (the number of samples in the demo).
- Price: the commercial terms of the exchange.

Each of these dimensions corresponds directly to a term a human participant would otherwise settle by hand. Data characteristics and sample counts map to the asset description and the usage scope, expressible as an ODRL policy, while price maps to the commercial clause of an eventual data-sharing agreement. The agents converge on the terms; the conclusion of any legally binding contract remains with the participating parties. The demonstration therefore realizes end-to-end, agent-to-agent negotiation of trade terms in a data-marketplace setting: ESJ supplies the marketplace and the catalog of tradable sensor data, and NEC's Automated Negotiation AI drives the agents that reach agreement. It instantiates the negotiation-agent role (Section 4.2) and the automated-negotiation capability (Section 4.3), and realizes the Agent-First interaction pattern (Section 2.1) operating over real IoT data. It is presented as a prototype use case in the context of the xIPF Consortium's activities.

The prototype was exhibited on 21 May 2026 at the xIPF Consortium's founding commemorative event, to an audience drawn from governmental, academic and business organizations. The prototype is a use case developed by NEC and EverySense Japan and is introduced here as such; it is not an official deliverable of the xIPF Consortium.

The significance of the prototype is that it shows the applicability of automation to the critical step of reaching agreement on the conditions of a data sharing. In doing so, it shows, in miniature, how an AI-native data space could scale the number of data-sharing relationships an organization can sustain far beyond what manual contracting allows, while keeping each negotiated term explicit and auditable.

5.4 PLIADES: AI-enabled integration of data lifecycles across data spaces

This pilot[24] context addresses a recurring problem across data spaces: even where the technical means to share data exist, the data itself is often too large, too poorly described, or too loosely governed to be used effectively by AI applications across organizational boundaries. Raw sensor and operational data accumulate faster than it can be sustainably stored; data assets are hard to discover or match to a consumer's needs without rich metadata; and participants are often reluctant to share data at all without stronger guarantees over who retains control of it once shared.

[24] PLIADES project. https://www.pliades-project.eu

The approach taken here is to apply AI not just to applications built on top of a data space, but within the data space's own lifecycle. For example, compression, filtering and normalization techniques are used to make data creation and storage more sustainable, reducing the footprint of data before it is ever shared. AI-based brokers and connectors, built on richer metadata, are used to improve discovery, helping a consumer in one domain find and match relevant data assets from another, in much the same spirit as the catalogue-search pilots described above, but extended across data spaces rather than within a single catalogue. And decentralized protocols are used to preserve the sovereignty of the organizations and citizens generating the data, so that improved discovery and sharing do not come at the cost of participant control.

The intended benefits are illustrated through examples spanning several sectors: in mobility, the aim is to supply the continuous, high-volume sensor data that automated and assisted-driving AI systems depend on; in healthcare and industrial settings, to support AI-driven human–robot interaction in ways that remain trustworthy when systems are acting on shared data; and in energy and the green deal, to let AI work as an optimization layer across data drawn from multiple organizations toward shared goals such as reduced carbon footprint. In each case, the same underlying lifecycle tooling, sustainable creation, discovery, brokering and sovereignty-preserving sharing is what is intended to make the sector-specific AI application viable at all.

As with the other pilots in this section, the work sits at the intersection of Data Spaces for AI and AI for Data Spaces: AI is used to operate the data space itself, while the data space in turn is shaped to supply the kind of data that downstream AI and robotics applications require.

5.5 Mission-KI Compliance Monitor

Mission-KI Compliance Monitor explores AI for data spaces in the governance layer[25]. Its aim is to generate machine-testable policies from contractual terms and to extend the Dataspace Protocol with AI-based compliance evaluation, so an agent can check a transaction against the rules that govern it. The proof of concept worked, but it also surfaced the limits of the approach. For legal evaluation, where reliability is the decisive property, classic deterministic policies remain more reproducible than AI-generated ones. Unit costs were also high, because professional distillation of the legal source material was outside the pilot's scope. The finding reinforces the separation set out in Section 4.4, where AI proposes and a deterministic layer decides.

5.6 Mission-KI Data Set Search Engine

Mission-KI Data Set Search Engine works in both directions at once, improving the data space for AI and using AI to improve the data space[26]. It adds an explicit data-quality dimension to catalogue assets through a connector extension that runs empirical data-science analysis, including anomaly detection, over a dataset and writes quality assurances into its metadata. This helps an AI system judge whether a dataset is fit for its purpose before requesting access. A second strand applies vector and similarity search over the catalogue, indexing entries by meaning so discovery extends beyond exact keywords. The proof of concept used a hash-based similarity search rather than an LLM, which keeps the method reproducible

[25] Mission-KI, Compliance Monitor. https://mission-ki.de/en/data-spaces/compliance-monitor
[26] Mission-KI, Data Set Search Engine. https://mission-ki.de/en/data-spaces/data-set-search-engine

and low cost[27]. Together the two strands deliver the empirical data-quality assessment and the semantic discovery that Section 2.1 describes on the findability and interoperability dimensions.

5.7 RoX: Data ecosystem for AI-based robotics

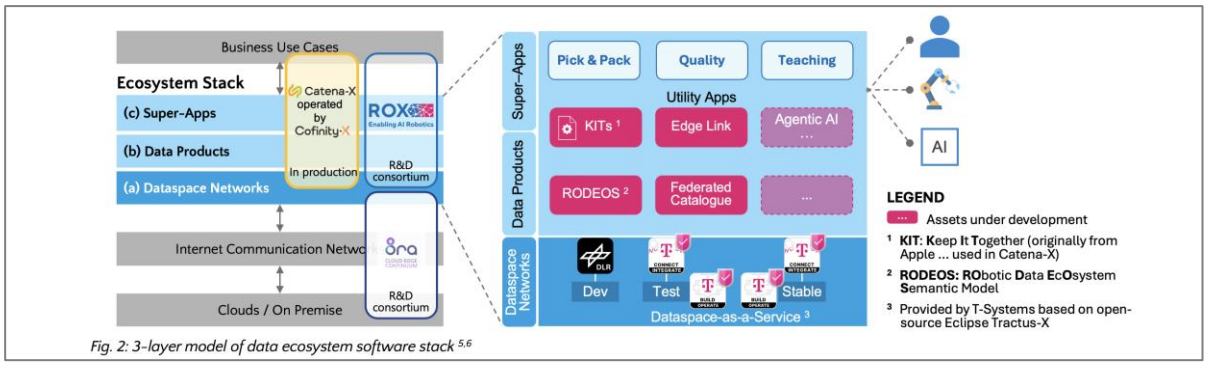

Figure 4: RoX data ecosystem stack and development

RoX (2024 to 2027) is a German consortium building a data ecosystem for AI-based robotics, supported by the Federal Ministry for Economic Affairs and Climate Action. It is a pre-competitive initiative in which robotics companies and research institutes that otherwise compete, among them DFKI and DLR, collaborate on shared foundations, reusable building blocks and standards. The problem it addresses is specific. AI can make robots more flexible, easier to deploy and more productive across manufacturing, logistics and services, but AI-based robotics depends on trusted access to machine, product, process, quality and supply-chain data that is scattered across companies, systems and lifecycle stages. RoX uses data space technology as the foundation for secure, governed and sovereign exchange of that data.

The operational infrastructure, provided by T-Systems, is a data space built on the open-source Tractus-X core and aligned with IDSA and Gaia-X. It runs in two environments, one for development, integration and testing and one as a stable environment for reliable consortium operation, which keeps experimentation separate from the demonstrators that have to keep working. On this foundation the consortium has shown live robotic cells for dynamic pick-and-place, quality inspection and teach-and-assemble, a first data-ecosystem application that visualizes data flows across participants, and federated catalogue capabilities that let applications discover and consume data from across the ecosystem rather than from a single source. A simulation application for pick-and-pack draws on live data assets delivered through the data space to help close the gap between simulation and real-world operation, a concrete case of physical AI built on governed data. Figure 3 shows the system stack and the elements still under development.

The outlook for RoX is agentic. The same governed data supply that feeds the robotic applications is intended to feed an agentic layer that automates and orchestrates tasks and connects agents to business processes under the data space's existing controls. RoX is one of the clearer demonstrations of the argument this paper makes: a data space turns distributed

[27] Mission-KI, Medas Verify Together App. https://github.com/Mission-KI/Medas-Verify-Together-App

industrial data into governed data supply for AI, and trusted, sovereign and interoperable data sharing is already becoming a practical foundation for AI at scale.

6 Outlook

6.1 What is clear today

Data spaces and AI are complementary rather than separate agendas: data spaces give AI a governed, high-quality data foundation, and AI gives data spaces the automation they need to operate at scale. The mechanisms that make this work already exist. The Dataspace Protocol, verifiable identity, machine-readable policies and provenance address the core trust and governance challenges that AI raises across organizational boundaries, and they do so without requiring a new protocol built specifically for AI. Agentic participation can be made trustworthy on these same foundations: an agent acts under a delegated identity that binds it to an accountable participant, governed by the policies, credentials and audit trails that already secure human participation. The pilots in Section 5 show this working in practice, with examples other organizations can reuse and adapt.

6.2 What is not yet clear

As participation becomes automated, it is not yet clear how data space organizations and their governance will evolve when most activity is initiated by agents rather than people, or how connectors and infrastructure will behave under the volume and speed of agent-to-agent traffic. Several mechanisms this paper describes are not yet standardized: how an agent's delegation is declared, how a model and its tests are published as governed assets and how cascading delegation is governed when one agent tasks another across organizations. The economics are unsettled too, since early work shows that AI-generated compliance checks can be less reproducible and more costly than deterministic ones, which bounds where AI can responsibly sit. And it remains to be seen how sector-specific trust sandboxes, such as a qualified body for health, are governed and held accountable over time. These are the questions the roadmap below begins to address.

6.3 A future roadmap

The path from today's foundations to data spaces that AI agents operate at scale runs through three broad stages. They overlap, and different sectors will move through them at different speeds, but the sequence is consistent.

The first stage is consolidation of what already exists. The Dataspace Protocol, verifiable identity, machine-readable policies and the trust hierarchy are in place, so the immediate work is to bind agent participation to them: the Delegated Agent Participant ID, agent discovery through the catalogue and the separation of probabilistic proposal from deterministic enforcement. The pilots in Section 5 are how this stage is validated and de-risked.

The second stage is standardization of the agent-specific layer. The gaps this paper identifies, an agreed way to declare agent delegation, to publish a model and its tests as governed assets and to govern cascading delegation across organizations, need to become shared specifications rather than per-project conventions. In parallel, the open questions above have to be answered in practice: how connectors behave under high-frequency agent-to-agent traffic, and how sector-specific trust sandboxes are governed over time.

The third stage is interoperability at scale. Agents operate across many data spaces as one connected infrastructure, the longer-term control plane described earlier, with the testbed

of Section 6.4 validating safety, compliance and cross-border interoperability before capabilities reach production. Sector bodies mature into qualified authorities for their domains, and a data space becomes the default way trustworthy AI consumes and serves governed data.

6.4 Concept of a "Data Spaces and AI Testbed"

As data spaces are equipped with a set of standardized mechanisms, including catalogues, connectors, identity mechanisms, access and usage policies, contractual rules, semantic models, protocols, and governance bodies, it is expected that both AI for Data Spaces and Data Spaces for AI will function appropriately within such an environment.

At the same time, AI systems operate at speeds far exceeding those of humans, continuously acquire new capabilities, and may occasionally make errors or exhibit behaviors that cannot be fully anticipated by their designers. For this reason, it is necessary to verify whether AI for Data Spaces and Data Spaces for AI can function reliably in practice and be operated safely and sustainably over time.

In addition, governance rules and regulatory frameworks governing the use of AI and data differ across countries and regions. It is therefore necessary to establish mechanisms that enable AI for Data Spaces and Data Spaces for AI to comply with these requirements while ensuring international interoperability.

Against this background, the IDSA envisions an international testbed project for data spaces and AI and will initiate efforts to test AI for Data Spaces and Data Spaces for AI through collaborative engagement among its members.

The concept of a "Data Spaces and AI Testbed" is not a standalone technology exercise. It can be positioned as a practical contribution to wider European and international policy priorities: scaling access to high-quality data for AI, connecting data spaces with AI ecosystems, simplifying trustworthy data use, and testing AI-enabled services under real governance conditions.

- **The EU Data Union Strategy (The Infrastructure Mandate):** This strategy acts as the architectural blueprint for the testbed. Its primary objective is to shift the European focus "from rules to results" by solving a critical structural vulnerability: the acute scarcity of high-quality, interoperable data required to train advanced AI models. It fulfills this by conceptualizing and deploying Data Labs, which serve as the critical technical conduits bridging the gap between static data repositories (Data Spaces) and dynamic, resource-heavy AI supercomputing.
- **The EU Apply AI Strategy (The Sectoral Mandate):** While the Data Union Strategy builds the pipeline to *access* data, the Apply AI Strategy ensures the resulting AI models are actively and safely *deployed* into the real economy. It dictates that the testbed's infrastructure (specifically the Data Labs) must be deeply tailored to specific priority domains such as healthcare, manufacturing, and scientific research, recognizing that legal constraints and metadata standards vary wildly across industries.

Together, these strategies drive the creation of the testbed to ensure Europe can transition from setting theoretical legislative standards to deploying tangible, sovereign AI capabilities.

The Core Concept: A Four-Pillar Pipeline

Driven by these strategic mandates, the Data Spaces and AI Testbed functions as an end-to-end, cyclical workflow that safely transforms raw, siloed data into highly competitive, market-ready AI models. It operates across four specialized layers:

- **Sourcing (Common European Data Spaces):** Decentralized, trusted environments where sector-specific data is securely stored and shared on a voluntary basis without relinquishing data sovereignty.
- **Processing & Preparation (Data Labs): The** "last mile" intermediary. Data Labs federate the raw data without extracting it, applying advanced curation, pseudonymization, and synthetic data generation. They provide Secure Processing Environments (SPEs) where competitors can pool data to train models without exposing trade secrets.
- **Model Training (AI Fabrics / Factories):** The world-class supercomputing infrastructure (backed by the EuroHPC JU) that ingests the standardized, "AI-ready" datasets directly from the Data Labs to execute massive foundational model training.
- **Validation (Testing and Experimentation Facilities - TEFs):** The real-world physical and digital sandboxes. Before commercial deployment, TEFs stress-test the trained AI models for algorithmic bias, safety, and operational robustness to generate evidence for compliance with applicable European regulations.

The Interaction Flow (The Testbed Concept)

When these four layers are connected, they form a continuous AI development assembly line:

- **Sourcing:** An AI developer locates relevant raw data residing securely within a data space.
- **Preparation:** The data is routed through a data lab, which cleans, anonymizes, and places it into an antitrust-compliant Secure Processing Environment.
- **Training:** The standardized, "AI-ready" data is immediately ingested by the supercomputing nodes of an AI fabric to train a foundational model.
- **Validation:** The trained model is deployed into a TEF, where it is benchmarked in real-world scenarios. If the TEF identifies operational flaws or algorithmic drift, the model is routed back to the beginning of the pipeline for continuous data stewardship and retraining.

6.5 What is next for the task force and the IDSA community

The task force will carry the treatment of agent identity, delegation and trust developed here into the IDSA Rulebook, shaping the AI Agents chapter and keeping it current as practice evolves.

Three lines of work follow. First, the testbed of Section 6.4 becomes a concrete international initiative, with members contributing data spaces, pilots and infrastructure and the collaboration patterns of Section 2 structuring what is tested. Second, the standardization gaps close, as the agent delegation credential, the publication of models as governed assets and the proposal-versus-enforcement separation move into the relevant standards and

working groups rather than staying project-specific conventions. Third, the evidence base grows with pilots and proofs of concept, each tagged to the interaction pattern and collaboration type it demonstrates, so the community learns from real deployments rather than theory.

The task force welcomes new contributors, pilots and engagement from data space initiatives, standards bodies and policy efforts, including the European Data Union and Apply AI strategies and their counterparts elsewhere. A companion paper on training data is already in preparation. During the work several issues to be tackled were identified and will be handled in the subsequent work of the task force. These are summarized in Annex A and maintained in full in the task force's GitHub repository. To take part, check the current backlog and join through the IDSA community.

Annex A: Task Force Backlog[28]

Outcome	Focus
DSP × AI protocols position paper	How the Dataspace Protocol relates to emerging AI protocols such as MCP and A2A, and why the two are complementary.
AI and data spaces reference architecture	A technical extension of the IDS Reference Architecture Model (IDS-RAM) showing where and how AI components fit into a data space.
AI governance model and usage control policies	The rules for AI in data spaces: when AI may use data, when it may not and how those rules are written and enforced, with granular policies across training, fine-tuning, inference and output reuse.
Trust framework and EU AI Act alignment	How data spaces deliver trust for AI and help meet regulatory obligations such as those of the EU AI Act.
Guidelines and best practices for AI integration	A practical guide for organizations integrating AI into their data spaces, covering technical, security, governance and enforcement aspects.
Compute-to-data and federated learning patterns	Technical patterns for running AI close to the data without moving it across organizational boundaries.
Agentic AI blueprint	How the Dataspace Protocol can be extended to support AI agents with proper identity, accountability and decision boundaries.
Metadata and catalogue extensions for AI	How data spaces describe, catalogue and quality-check AI assets, including models, datasets and services.
Use case collection	A curated set of real-world examples: what works, what does not and what is still needed.
International testbed for AI and data spaces	A shared international environment where members contribute data spaces, pilots and infrastructure to validate AI for Data Spaces and Data Spaces for AI in practice, building on the concept in Section 6.4.

Table 8: Backlog outcomes identified by the task force

[28] International Data Spaces Association, Task Force Data Spaces and AI, project backlog. https://github.com/International-Data-Spaces-Association/Task-Force-Data-Spaces-and-AI/blob/main/Backlog.MD

CONTACT

International Data Spaces Association

Emil-Figge-Str. 80
44227 Dortmund | Germany

phone: +49 231 70096 501
mail: info@internationaldataspaces.org

WWW.INTERNATIONALDATASPACES.ORG

international-data-spaces-association