INTERNATIONAL DATA SPACES ASSOCIATION

Position Paper | Version 1.0 | June 2025

Observability 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-Str.80
44227 Dortmund
Germany

Editor

Sebastian Steinbuss, IDSA

Cite as

Steinbuss S., Observability in Data Spaces, International Data Spaces Association, June 2025
https://doi.org/10.5281/zenodo.15647484

Authors & Contributors

Peter Koen, Microsoft
Ilknur Chulani, IDSA
Gabriele Bozzi, IMEC
Daniel Hommen, Orbiter
Maximilian Schönenberg, Orbiter
Philippe Calvez, Acatech
Julia Pampus, Fraunhofer, ISST
Arian Firouzbakhsh, IONOS
Markus Spiekermann, Huawei
Theo Dimitrakos, Huawei
Marko Turpeinen, 1001 Lakes
Luiza Brandao, Acatech
Olaf Gerd Gemein, Orbiter
Sourabh Bharti, MTU
Francisco Morais, CCG
Wouter Van den Berg, TNO
Sara El Gaily, 6G Health Institute
Andreas Mueller, Schaeffler
Petteri Kivimäki, NIIS
Antonio Pedro Salvado, Tice.pt
Werner Jost, T-Systems

Table of Content

List of Figures

List of Tables

1 Introduction

1.1 Motivation

Data spaces are meant to be a trusted system where participants can share, and exchange data under a controlled governance scheme. Even though data sharing happens always between the participants under some circumstances, such transactions need to be observed for different reasons. Observability of the transaction can be a requirement only between the parties involved or as a measure indicated by regulations or the governance scheme of the data space. **This document shall investigate the need for observability in data spaces, reflect the current findings on the subject, provide some insights into relevant use cases and propose potential solutions.**

The need for observability and requirements is stated in the IDSA Rulebook. Observability may result in the exchange of regulated or highly valuable datasets. Likewise, marketplaces or general billing and charging may require the observation of the transaction. Depending on the requirements, this could lead to a centralized or a decentralized solution to realize the observability of the transaction. The approach to identifying an appropriate solution depends on the requirements stated in the general governance framework of the data space, which is managed by the Data Space Governance Authority or the individual contract between the participants. Furthermore, it is important to consider for whom the collected information will serve after it is collected. Different kinds of third parties may need to be enabled to read or analyze the collected information, e.g., a third-party auditor, a governmental authority, or any third party identified in case of incidents. Such information needs to be provided or evaluated before the transaction, during the transaction, or even after the transaction.

In this regard, observability in data spaces enhances trust by attesting that data is used as agreed. It aligns with data spaces' governance principles and strengthens compliance monitoring and, therefore, protects data providers' and consumers' interests. Observability imposes some technical complexity of tracking data usage and has privacy implications for monitoring. Legal constraints across jurisdictions, potential resource-intensive implementation, and subsequently a potential impact on data processing performance need to be addressed.

1.2 Why Do We Need Observability

Observability of activities in a data space can serve many business purposes. It can be the notarization of agreed data-sharing contracts between two parties by an independent 3rd party, proof of existence of data-sharing contracts for required regulatory reporting, auditing of contract policies and claims, business processes like billing at a marketplace, enabling

continuous verification and attestation that data usage (access, processing, and transfer activities) comply with contractual obligations and governance frameworks and many more[1].

Observability may be a function implemented solely by the Data Provider and/or the Data Consumer, or alternatively, it may involve an independent 3rd party, for instance, a notary service or a regulatory body, depending on the use case requirements. Data space governance framework may mandate observability processes by default for certain uses, types of data or data providers, consumers.

Observability data might be as sensitive as the data shared under the data-sharing contract as it might divulge important information about business processes and connections between participants of a data space to third parties. Analyzing the observability data of an entire data space might yield detailed information about the business activities in the data space and thus reveal sensitive information. It is, therefore, of utmost importance to establish trust not just between the two parties sharing data but also any 3rd party that may be receiving observability data.

The easiest way to achieve a trusted relationship with an observer is to treat the observer as just another participant in the same data space and observability data as just another data stream that requires sharing. This enables the reuse of existing trust creation mechanisms, as well as the establishment of policies governing the permitted use of observed activities.

Any data space can contain any number of observers, depending on use cases, domains, and needs of the participants. Regulations, contracts, or data space rulebooks may provide guidance in this matter. As Observers act like any other participant negotiating data-sharing contracts, they are bound by the same rules as all other participants. Observability, thus, is just a service any participant can provide to any other participant. An observer is a business process role and not a technical function. At the technical layer of a data space, observability does not require any special implementations.

However, due to the nature of observability data and the potential need for special treatment of its semantic models and policies, it is very likely that dedicated Data Plane Extensions for observability will be developed, potentially based on ODRL and ODRL extensions and profiles. Nevertheless, observability may also impact the Control Planes to a certain extent. Standardizing and sharing the responsibility for developing Observability Data Plane Extensions will highly increase the reliability of a data space while reducing the operational costs for its participants.

[1] See https://docs.internationaldataspaces.org/ids-knowledgebase/idsa-rulebook/idsa-rulebook/3_functional_requirements#observability for more information on the functional requirements of observability.

1.3 Relationship with Data Provenance and Traceability

Some use cases need additional data (metadata) over the actual data being shared for the purposes of auditing and compliance. Depending on the use case, it might be necessary to have a trace of transactions taking place in data space or to know who has had access to certain data.

The requirement for observability, traceability, and provenance tracking is usually found in highly regulated industries or in cases dealing with high-value data. Furthermore, in increasingly digital socio-economic ecosystems, data exchange observability requirements operate across interconnected regulatory layers, each addressing specific aspects of data-sharing governance. At the foundational level, cross-sector regulations like the Digital Markets Act (DMA)[2] and Free Flow of Non-Personal Data Regulation[3] establish baseline requirements for data exchange monitoring across all industries. These are complemented by sector-specific regulations, such as MiFID II[4] in finance, REMIT[5] in energy, and HIPAA[6] in healthcare, which impose additional observability requirements tailored to industry-specific risks and compliance needs. A third layer consists of competition-related controls, particularly critical in scenarios involving data exchanges between market competitors or within joint ventures, where observability serves to prevent anti-competitive practices and ensure market fairness. These regulatory layers interact dynamically, creating a complex framework where organizations must implement comprehensive observability mechanisms that satisfy both horizontal (cross-sector) and vertical (industry-specific) requirements while maintaining appropriate security controls, audit trails, and compliance documentation. This multi-layered approach ensures that data exchanges are not only compliant with general data protection principles but also meet the specific observability requirements of their operational context and competitive environment.

[2] https://digital-markets-act.ec.europa.eu/index_en
[3] https://digital-strategy.ec.europa.eu/en/policies/non-personal-data
[4] https://eur-lex.europa.eu/eli/dir/2014/65/oj/eng
[5] https://www.acer.europa.eu/remit/about-remit
[6] https://www.ncbi.nlm.nih.gov/books/NBK500019/

It is important to distinguish between transactions between data space participants during the control phase, i.e. regarding the data-sharing contracts, and the actual data-sharing phase as shown in the Figure below.

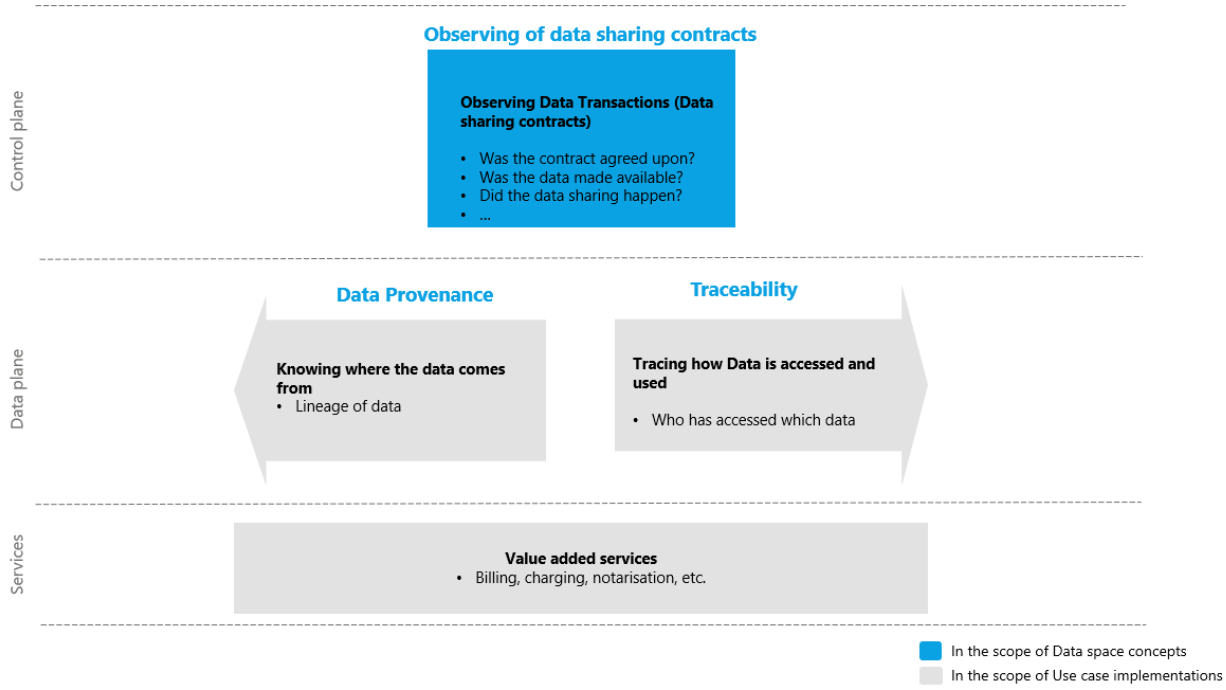

Figure 1 Distinguishing observability of data-sharing contracts from data provenance and traceability

Observability, provenance, and traceability are important measures to achieve trust in data spaces.

Observability, as described in this document, is a tool to enable monitoring and verification of data-sharing contracts throughout their lifecycle[7]. It enables support for enforcement of data-sharing contracts and regulatory and contractual compliance.

Data Provenance tracking complements this by documenting the lineage, transformation, and utilization of the actual data being shared. **Both data provenance and traceability are linked to the observability in data spaces,** and could be required for regulatory and contractual compliance, **however, they are not in the scope of this document.**

[7] The lifecycle of data sharing contracts needs to be further defined in the IDSA Rulebook.

In this sense, observability of data-sharing contracts, Data Provenance, and Data Traceability are not limited to technical implementations but require processes and rules as part of the Data Governance by a data space participant and governance rules for participation management by the Data Space Governance Authority to achieve compliance with regulation and contracts.

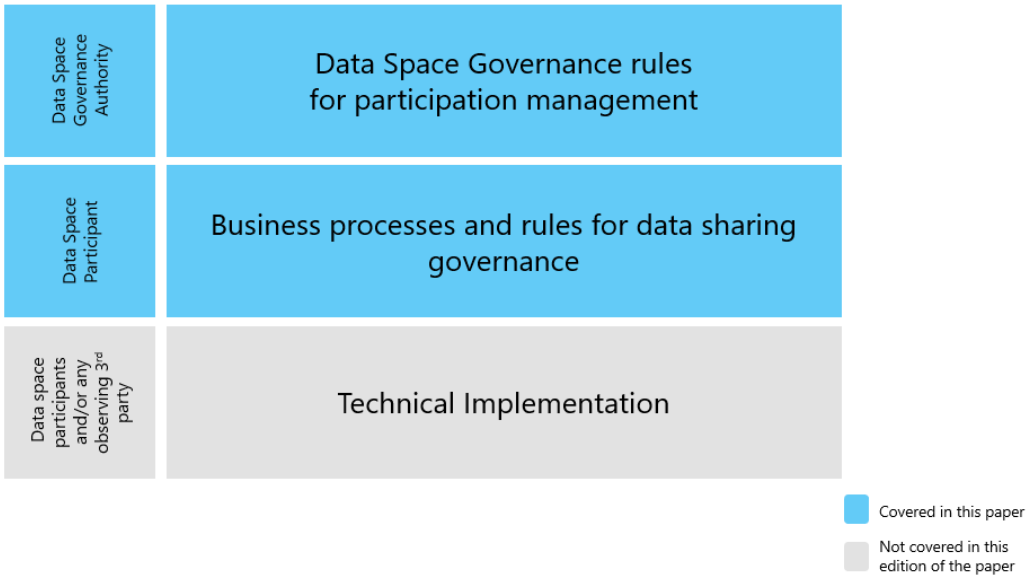

Figure 2 More than technical elements are needed to enable observability, Data Provenance and Traceability

1.4 Evolution in Data Spaces: From the Concept Clearing House to Observability

The Clearing House concept introduced in IDS RAM 4 was the first in the context of data spaces to tackle something remotely close to what is now discussed under the topic of observability.

As a general term, a clearing house is a financial intermediary that facilitates the clearing and settlement of transactions in financial markets. Its primary role is to ensure that trades between buyers and sellers are completed smoothly, efficiently, and securely. Clearing houses are commonly associated with stock exchanges, derivatives markets, and other financial markets.

The Clearing House role and component in RAM 4.0 had multiple functionalities. Its main purpose was clearing and settlement services for financial services, but it also logged all relevant information and provided means for provenance tracking and usage control.

One problem is that the term Clearing House is quite overloaded in the context of data spaces. For instance, the Gaia-X Digital Clearing House serves a completely different purpose as part of the Gaia-X trust framework, having no relation with the concept of observability as described in this document.

Another key point is that the settlement and clearing services are no longer envisioned as a core functionality of data spaces in the context of the IDS RAM and the IDSA Rulebook. Even though a Clearing House could facilitate and support certain transactions in a data space, this is no longer a functionality defined by IDSA. However, it could still be implemented as a value-adding service as part of a use case.

For these purposes, the Clearing House role and component are considered deprecated for the next version of the IDS Reference Architecture Model. It will be replaced by an observer role, as RAM 5 will focus more on the observability aspect, as explored in the rest of this document.

2 Core Concepts of Observability in Data Spaces

A data space is seen as the mechanism to negotiate trust for data sharing and to agree on data-sharing contracts with the actual data flow happening on private channels. Hence, it is important to distinguish between activities that can be observed within a data space and those that lie outside of the scope of data space observability but might still be important for the end-to-end trusted data sharing.

Also, it is important to distinguish between the observability of data space activities versus regular IT Operations telemetry. While both are important in an end-to-end solution, only the observability of data space can be defined generally through this architecture model. The observability of service telemetry is highly dependent on the specific implementation of the individual component and can rely on or impose a business model of a data space or its participants. However, the principle of sharing observed telemetry through a data-sharing contract in a data space applies to this data as well.

2.1 Types of Observability

In this document, we distinguish the following types of observability in data spaces:

- Observation of Dataspace Protocol activities & states
- Observation of Service Telemetry.

2.2 Observation of Dataspace Protocol Activities & States

In this category, any states and state transitions of the Dataspace Protocol can be observed. For this purpose, the connector or its related services needs to keep log entries (which need to be defined) of state transition requests and successes and failures of those state transitions for any state machine of the Dataspace Protocol (see https://github.com/eclipse-dataspace-protocol-base/DataspaceProtocol/blob/main/specifications/negotiation/contract.negotiation.protocol.md for an example of a state machine of the Dataspace Protocol).

The attributes of messages sent to the connector endpoint must provide comprehensive, machine-readable information structured according to standardized schemas to enable unambiguous process identification (and could rely on W3C standards such as, but not limited to, PROV-O[8] for provenance tracking, ODRL[9] for usage policies, ...), and the two participants involved in the data-sharing contract. This allows for an implicit creation of namespaces to segment the log entries per participant or even per negotiation and to correlate those in future analysis.

The three state machines[10] of the Dataspace Protocol cover the following three areas of observability: Catalog, Contract Negotiation, and Transfer Process, namely. The following tables depict what can be observed in each state:

Table 1: Observation of the catalog states

Catalog		
State	**Participant**	**What can be observed**
Catalog Request	Provider	Which participant requested the Catalog, and which optional filters and authorization tokens have been provided?
	Consumer	If successful, the ID of the resulting Catalog.
Dataset Request	Provider	Which dataset has been requested by the consumer, and which authorization token was provided?
	Consumer	if successful, the address of a valid instance of a Dataset.
Catalog Error	Both	If a Catalog Message resulted in an Error, this will provide the implementation specific error code and optionally a set of reasons why the preceding operation failed.

[8] https://www.w3.org/TR/prov-o/
[9] https://www.w3.org/TR/odrl-model/
[10] The Catalog specification does not include a state machine. The state machine for the Negotiation is described in detail in this section https://github.com/eclipse-dataspace-protocol-base/DataspaceProtocol/blob/main/specifications/negotiation/contract.negotiation.protocol.md#state-machine, the Transfer Process state machine is described in this section https://github.com/eclipse-dataspace-protocol-base/DataspaceProtocol/blob/main/specifications/transfer/transfer.process.protocol.md#state-machine .

Table 2: Observation of the contract negotiation states

Contract Negotiation		
State	**Participant**	**What can be observed**
Contract Negotiation Request	Provider	Which Contract Negotiation (CN) did Consumer (C) request? A new or an existing one? Was the CN successfully found, and is C authorized to access it? If it is a new CN, what are the terms of the proposed CN?
	Consumer	Has a Contract Negotiation been initiated by posting an initial offer? What are the terms of the proposed CN?
Contract Offer	Provider or consumer	Can make offers to the CN process. Once the offer is accepted, an acceptance event will be logged.
Contract Acceptance, Agreement and Verification	Both	The accepted terms of the CN will be logged by both parties and the agreement will be logged as well as a verification of the agreement.
Finalized	Both	The finalization of the Agreement is logged
Terminate	Both	Should any of the steps above result in an error, then the CN will be terminated, and the error will be logged. This can have numerous reasons, e.g., State Transition Errors (wrong attributes, not reconcilable terms...), requests to non-existent CNs, unauthorized access, ...

Table 3 Observation of the transfer process states

Transfer Process		
State	**Participant**	**What can be observed**
Request	Both	Which dataset has been requested under which data-sharing agreement by which consumer.
Start	Both	Once a dataset is available for access by C or P has started pushing data to an endpoint provided by C a log entry will indicate the start of the data transfer.
Completed	Both	If the transfer has been completed a log entry will indicate that P has finished transferring data.
Suspended	Both	When C or P suspends the transfer the log entry can contain the reason why the transfer had to be suspended.
Terminated	Both	Should any of the steps above result in an error, then the CN will be terminated, and the error will be logged. This can have

		numerous reasons, e.g., State Transition Errors (wrong attributes, not reconcilable terms…), requests to non-existent CNs, unauthorized access…

Please note that a Transfer Process (TP) involves two parties, the provider and the consumer. However, data might not leave the location where it is stored if the data plane is implemented with a code-to-data mechanism. Also, there might be myriads of different transfer mechanisms, depending on the data planes involved (streaming data, one time transfer of blobs, database access, etc.…). The state machine of the TP therefore doesn't include the data technology specific sharing details. Rather it provides messages to orchestrate the high-level operation of the data plane. Data technology specific logging of sharing events is highly dependent on the implementation detail of each technology and therefore not defined here, but something that needs to be handled custom per data plane.

Observing the messages and states of the Dataspace Protocol can provide rich insights into the trusted data-sharing processes within a data space. Easy correlation between provider and consumer participants allows for a clear understanding of what contracts for which data have been negotiated and executed by whom.

2.3 Observation of Service Telemetry

End-to-End implementations will also require additional telemetry like service uptimes, performance data, and other measurements of the solution. However, as those are implementation specific, they cannot be detailed in this architecture model for data space observability and need to be agreed upon as an operational aspect of a specific data space.

3 Technical Implementation

3.1 Observability at the Technical Layer

At the technical layer of a data space, the collection of observability data involves both standardized requirements and implementation flexibility. It is recommended that a Data Space Governance Authority (DSGA) establishes a common schema for describing core observability data related to fundamental data space services. This standardized schema ensures interoperability and consistent interpretation of observability data across the data space.

Each participant remains free to implement, collect, and store observability data using whatever formats and technologies best suit their infrastructure and operational requirements. They can transform this data to conform with the standardized schema when sharing it with other data space participants. This approach balances technological autonomy with interoperability needs. Observability should not be limited to connectors alone but should encompass all components and services that participate in data space operations, including catalog services, identity providers, policy enforcement points, and any other

services that contribute to the data space ecosystem. Each of these components may generate valuable observability data through their specific tooling that contributes to a comprehensive view of data space health and operations.

Generally, observability data can be shared within a data space with other participants like any other data. Participants maintain decision-making authority over which observability data they share, with whom, and under what conditions. However, certain data contracts or regulatory requirements enforced by the DSGA may mandate the sharing of specific observability datasets. When the DSGA enforces mandatory observability requirements, these are typically limited to specific agreed-upon cases within the data space or, more commonly, to satisfy external audit requirements necessitated by applicable laws and policies. All participants must have the capability and agency to share the required observability data as specified in negotiated data contracts.

It is recommended that the DSGA establishes a standardized data plane (or multiple planes) for the sharing of observability data within the data space to facilitate interoperable observability.

3.2 Semantic Models of Observability

A standardized semantic model for the sharing of log entries originated by messages on the Dataspace Protocol (and potentially other protocols like the Decentralized Claims Protocol) will greatly simplify the implementation of observability across a multitude of data spaces and thus be a worthwhile open-source specification project and/or standardization project. This will enable the development of a rich ecosystem of value-added observation services which can be provided to a multitude of data spaces, as well as the development of tools for reporting, analyzing and auditing of data spaces activities.

At the time of writing, the authors are not aware of any projects for the definition of Semantic Models for observability. As future work, this document may be extended based on contributions from the data space community.

4 Business Process Integration

As observability is just another data-sharing contract at the technical level, it is up to the business process on how to implement observability within a data space. Depending on the design of the data space, the DSGA might require one central observer service to receive all observations, a federation of observers to jointly ensure observability, potentially segmenting observation services by domain or jurisdiction, or even a fully decentralized approach. However, it is also possible to have an open market of observation services providing value-added services to other participants of the data space like notary services, data accounting/payment processing, dispute resolution, proof of execution, etc.

4.1 Exemplary Process for Setting Up Observability

The detailed process might vary depending on the rules and requirements of the data space and the implementation of the connector components. However, a generic flow can be described as follows, just as an example:

1. Participant C (Consumer) requests a Contract Negotiation CN from Participant P (Provider)
2. Contract Offer CO from P contains a term that requires the use of an observation service to observe the CN and/or the Transfer Process TP
3. P includes information on acceptable Observers O(P) within the CO
4. P includes claims and evidence about established Data-Sharing Contracts for Observability (DSCOs) with the suggested O(P) in the contract offer.

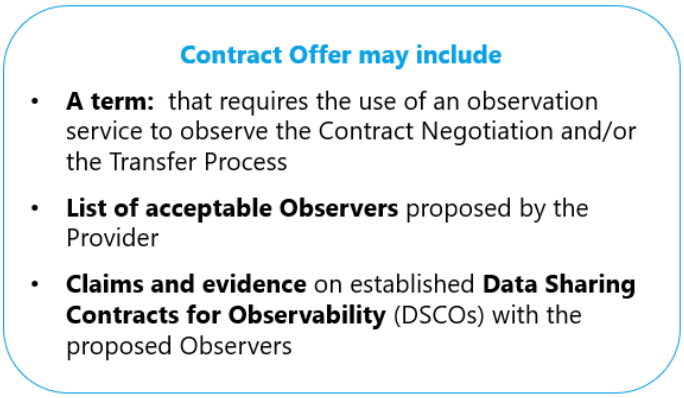

Figure 3 Observability related elements that may be included in a contract offer

5. C matches the list of O(P) against its internal list of available DSCOs.
   5.1. If a match is found and is deemed suitable for the requirements of the use case and the rules of data sharing, the matching DSCO claim and evidence are provided in the CN process to P
   5.2. If no match is found, C selects one (or multiple Observer from O(P) and requests a CN for a DSCO. Once one (or multiple) DSCOs are negotiated, C proceeds with the process of 5.1
6. C and P verify with the selected Observer O the existence of the DSCO of the other party.
7. C and P finalize the negotiation of their CN. The negotiation process results in the generation of log entries on both sides. Which get shared through the established DSCOs with the Observer.
8. Finalization of CN negotiations is logged with O. Depending on the agreed terms, the process might stop here or continue to observability of the TP.
9. It is advisable to recheck the existence of the DSCOs at the beginning of the TP whether they are still active, and whether all parties involved are available.
10. Log entries from the TP will be shared through the data streams agreed upon in the DSCOs as during the negotiation phase.

Note that it is possible that multiple Observer and multiple DSCOs have been negotiated, e.g., separated by phase of the data-sharing process (negotiation vs execution), control plane vs data plane observation (data space activities vs data technology telemetry). Also note that while connectors' and data planes' components might collect extensive telemetry data they might filter and transform this data according to the use case and confidentiality requirements.

While this general process is applicable in many situations it is expected that individual implementations will vary. For instance, a DSGA could require the setup of specific DSCOs at the time of joining a data space.

Note that it is also possible for C or P to act as O without the need for a third party. By signing the log information, sharing it and comparing it to its own log entries, any one of the two participants can confirm an agreement on the logged information[11] and regular cryptographic signing mechanisms can ensure tamper proof storage of the observed transaction. This would allow full confidentiality between two participants with the option of proofing observed data sharing to an auditor or regulator, if necessary, without involving any third-party observer.

Figure 4 Sequence diagram for an exemplary process for setting up observability of a data-sharing contract is included to help visualize this exemplary process for setting up observability of a data-sharing contract:

[11] If additional trust is needed, the logged information could also include a timestamp provided by a trusted Timestamping Authority (TSA) and/or evidence on the signing certificates' validity (e.g., OCSP response, CRL). All mechanism that are available for Trusted Data Sharing contracts are also available for Data Sharing contracts for Observability.

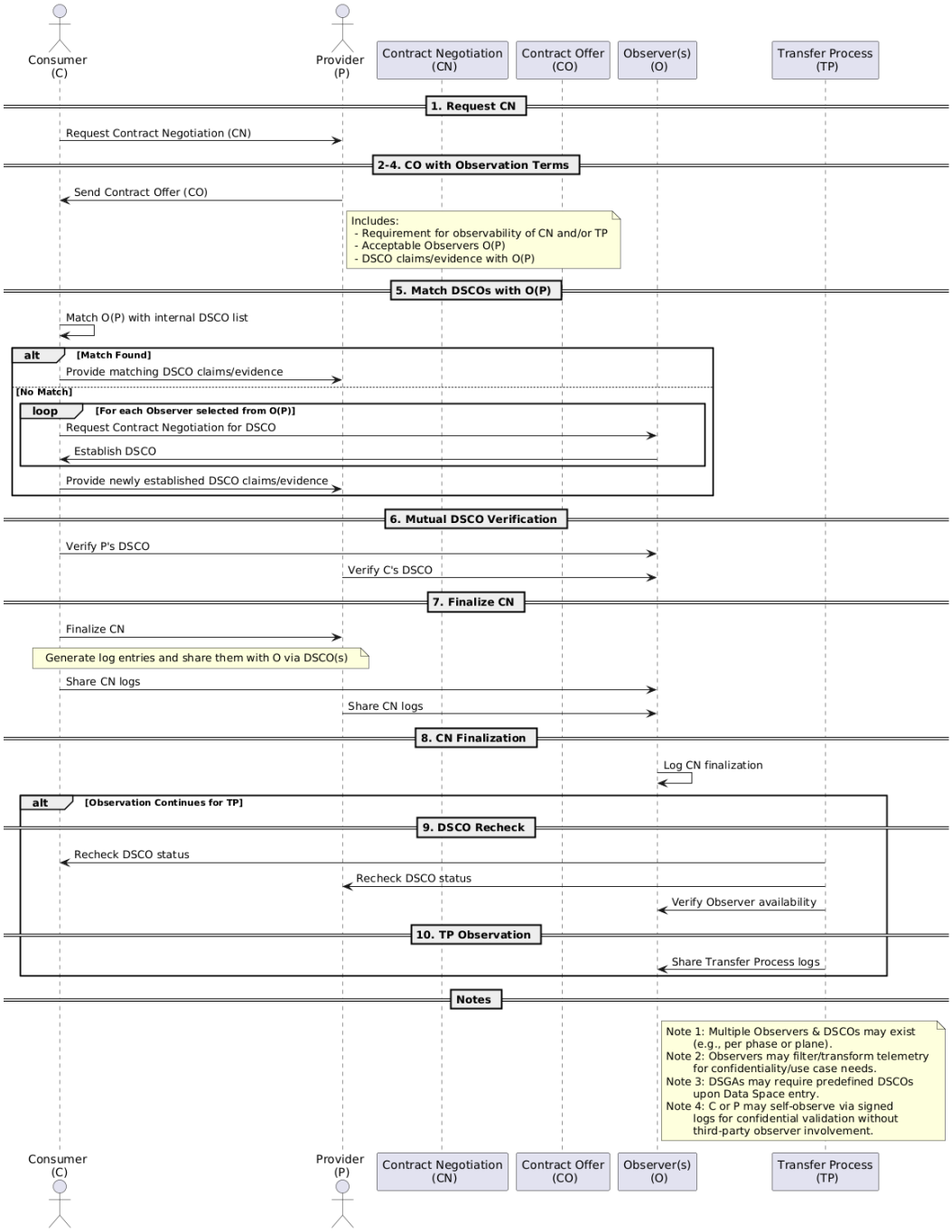

Figure 4 Sequence diagram for an exemplary process for setting up observability of a data-sharing contract[12]

[12] For more details, please refer to the Activity diagram in Annex-A.

5 Use Cases

This document provides an overview of observability in data spaces. The justification for the degrees of freedom in processes, rules, policies, standards, and technical implementations is rooted in the potentially different requirements of data spaces and their participants coming from regulation, contracts, business aspects, operational aspects, and technical needs. This section on use cases shall illustrate parts of such requirements and approaches.

5.1 Billing and Clearing of Transactions

Billing and clearing of transactions are one potential use case, which could be enabled by observability. Nevertheless, more specifications and implementation beyond observability could be required to implement such a case.

Billing and clearing transactions in a data space involve the exchange of data between providers and consumers, often governed by usage-based pricing models, access policies, and service agreements. Ensuring transparent, accountable, and error-free financial transactions require tracking who accessed what data, when, and under what terms. Observability can be crucial in monitoring these interactions, providing real-time insights into data exchange, usage patterns, and financial reconciliation.

For example, in a manufacturing data space, a company purchases machine performance data from a sensor provider. Observability monitors data flow, API request patterns, and access logs, ensuring that only authorized usage is billed while detecting unexpected overcharges or misuse. By embedding observability-driven monitoring, data spaces ensure fair, accurate, and compliant billing and clearing processes, reducing financial risks and strengthening trust between providers and consumers.

5.2 Regulatory Compliance of AI Model Training

AI thrives on vast datasets, including data from the creative industries, whose content fuels model training. However, much of this data is used without the consent of the original content creators, generating significant value while leaving the creators uncompensated. This imbalance underscores the need for data provenance, fair compensation, and transparent collaboration between AI developers and content owners. This is an issue often encountered at the Media sector and the Creative Industry, whose claims against big Generative AI players have already reached the courts[13].

Data spaces offer a transformative solution, ensuring content protection while enabling structured licensing models. This may allow content creators, who are the data providers, to be compensated for their assets fairly and provide AI developers with a trusted, compliant source of high-quality data. By enforcing observability, data spaces help track content origins,

[13] https://www.heise.de/news/Skip-the-links-Wall-Street-Journal-verklagt-KI-Firma-Perplexity-9989199.html

ownership, and usage conditions — ensuring adherence to regulations like the EU AI Act and reducing legal risks. More than just a safeguard, data spaces enable traceability, accountability, and equitable value distribution, bridging the gap between creative industries and AI companies. This model fosters a sustainable AI ecosystem where innovation and fairness coexist. By its turn, observability supports the guarantee that data is not unrightfully used to train LLMs without respecting copyrights or the usage conditions established by the data rights holders. It offers the possibility to enforce compliance and to make sure that AI is trained with data from providers that not just consent and are aware of this use but also are paid for it. Thus, observability might offer legal certainty for the AI players and addresses the remuneration issues behind the media and creative industries, who, in their turn, can provide quality and reliable data.

The Trusted European Media Data Space (TEMS)[14] includes a use case to explore the potential of offering quality data and media content to AI players, in a trustworthy and sovereignty environment. For that, media participants will be able to select the data they aim to offer for the AI training and LLMs, define the conditions, including the remuneration related ones, and the copyrights linked to their productions. Through the data space, AI players can find this data, already linked to the legal conditions and applicable compliance, and be able to contract with the data providers. Observability plays a special role in this context, once it allows the data providers to be certain of the use of their data, and of the enforcement of the polices defined by them and in the context of the data space. For the AI players, it represents authenticity of the source of the data and its quality and the possibility to fulfil obligations related to AI transparency and explicability. Further, it allows to establish the source of the data used for their AI models, adjudicating copyrights issues, and avoiding legal conflicts[15].

5.3 Contractual Compliance

Contractual compliance within a data-sharing environment can be supported through either designated observers or an integrated observability functionality. The choice depends on the specific requirements of the participants involved. To ensure appropriate compliance measures, contracts must explicitly define which participants or actors are involved in each use case and the roles they play.

Additionally, it is crucial to clarify the necessary legal agreements governing data sharing, usage rights, and liabilities across different use cases. All contractual arrangements must align with the overarching data space governance framework and comply with applicable regulations and laws.

Mobility Data Space use case

Rationale and context

[14] https://tems-dataspace.eu/
[15] https://techcrunch.com/2024/11/22/openai-accidentally-deleted-potential-evidence-in-ny-times-copyright-lawsuit/

In today's rapidly evolving mobility landscape, data is the cornerstone of innovation, driving the development of smarter, more efficient transportation solutions. The Mobility Data Space (MDS) is a pioneering initiative designed to foster collaboration among companies, organizations, and institutions. By bringing together entities that require data to create innovative mobility solutions with those that seek to monetize their data assets, the MDS establishes a dynamic marketplace where data exchange can thrive.

Context

The MDS serves as a vital platform, bridging the gap between data providers and data consumers in the mobility sector. It creates an ecosystem where diverse stakeholders—ranging from automotive manufacturers, public transportation operators, technology companies, to urban planners—can securely and efficiently share data. This collaboration not only fuels innovation but also enables the development of cutting-edge mobility solutions that address the needs of modern societies.

Why

One of the fundamental reasons for establishing the MDS within this framework is to ensure compliance with German cartel laws. The legal landscape in Germany, particularly in terms of antitrust regulations, mandates a careful balance between fostering collaboration and preventing anti-competitive practices. By operating within the guidelines of these laws, the MDS ensures that data sharing and collaboration occur in a manner that is fair, transparent, and legally sound. This compliance not only safeguards the participants from legal risks but also enhances the trust and credibility of the Mobility Data Space as a whole.

Who is responsible for observing

The Logging House of the Mobility Data Space has been designated as the primary entity responsible for monitoring the interactions within the MDS. This entity will diligently observe all contract agreements and data transfers between participants to ensure that the processes align with the established rules and guidelines of the Mobility Data Space.

What needs to be observed

The Logging House's role involves carefully tracking and logging the details of all contract agreements and data exchanges among the participants in the MDS. This includes monitoring the terms of the contracts, the data being exchanged, and the compliance of these exchanges with the legal framework governing the MDS. The transferred data itself is excluded from the logging process although an asset's metadata is logged. The goal is to ensure that all transactions are conducted transparently, securely, and in full accordance with the relevant regulations.

Who needs access to observability data

Once the observability data is collected, it may need to be accessed by specific stakeholders for various purposes. Firstly, the Mobility Data Space itself will require access to this data to generate key performance indicators (KPIs) and to provide support to participants when

necessary. Additionally, there may be instances where a neutral third party, such as the Bundeskartellamt (Federal Cartel Office), needs access to this observability data for regulatory reasons. This ensures that all data exchanges within the MDS are compliant with German cartel laws and that any potential anti-competitive behavior is promptly identified and addressed.

Table 4: Use of observability in the Mobility Data Space

Use of observability in the Mobility Data Space			
Step No.	**Step Name**	**Step Description**	**Step Actor**
1	Data Exchange Initiation	A data consumer requests access to a specific dataset from a data provider within the MDS, following contractual agreements and predefined conditions.	Data Consumer, Data Provider
2	Contract Logging	The Logging House records the contract details, including metadata about the agreed-upon data exchange, terms of use, and compliance requirements, ensuring transparency.	Logging House
3	Data Transfer Execution	The agreed-upon data exchange takes place between the data provider and consumer, while only the metadata of the transfer is logged for observability and compliance.	Data Provider, Data Consumer
4	Compliance Monitoring	The Logging House continuously monitors data transactions, verifying that exchanges comply with EU regulations and MDS guidelines, ensuring fair and legal data-sharing practices.	Logging House
5	KPI Generation & Performance Analysis	The Mobility Data Space analyzes observability logs to generate key performance indicators (KPIs) related to data exchanges, efficiency, and compliance adherence.	Mobility Data Space
6	Incident Detection & Anomaly Reporting	Observability data is analyzed to detect potential violations, anomalies, or anti-competitive behaviors. If issues are found, reports are generated for further investigation.	Logging House
7	Regulatory Audit & Third-Party Access	A government regulatory body (e.g., EU partner) may request access to observability data to assess compliance and investigate any potential legal concerns.	Regulatory Authority

8	Compliance Enforcement & Support	If non-compliance is detected, the MDS takes corrective actions, notifying relevant stakeholders and providing guidance on resolving legal or contractual issues.	Mobility Data Space, Logging House

5.4 Analyzing Observed Telemetry

Either the observability service provider or, in case none exists, any of the two sharing parties can easily correlate and verify observation logs and use a plethora of analysis methods to derive additional value from the observations. Such use cases of analysis might be:

- Cryptographically signed log entries can be used as a source of truth in the resolution of disputes.
- Assumptions about the state of the data-sharing process can be mutually verified
- Service providers can process log entries to provide mandatory regulatory reporting
- Payment processing for shared data can be automated based on verifying log entries of the TP from both participants
- Claims about data quality can be verified between participants

And many more are possible to enable a vibrant ecosystem of trusted data sharing!

5.5 Monitoring the Adoption of Standards

Motivation

Standards Development Organizations (SDOs) invest a lot of effort in developing and maintaining common data models but often lack empirical data about how these models are actually used in real-world data exchanges. This makes it difficult to make evidence-based decisions about model changes, deprecations, and extensions. Instead, SDOs gather indirect evidence in the form of users' feedback by holding consultation rounds or organizing workshops and discussion meetings with the user community. These are time-consuming and costly, and rely on active participation from users to voice their needs.

Relation to observability in data spaces

Observability in data spaces would provide SDOs with concrete evidence about the actual adoption and usage of their standards in the field.

While SDOs in some cases can currently measure the number of registered connectors claiming to support their data models, observability would add actual usage metrics and patterns. This provides a more complete picture of standard effectiveness. The use case demonstrates how observability data can be analyzed to improve the quality and practicality of data space standards themselves.

Example implementation in a data space

The use case could be implemented by having an Observer service that monitors the structure of exchanged data (without accessing actual data content). It then aggregates usage statistics across multiple data exchanges, which is then provided in anonymized form to SDOs to inform their standardization decisions.

For SDOs this has some concrete benefits:

1. Evidence-based decisions about model changes:
    - Which optional elements are frequently used vs rarely used
    - Which elements could be candidates for deprecation
    - Which cardinality constraints match real-world usage patterns
    - Which parts of the model generate the most validation errors
2. Prioritization of change requests based on usage intensity of affected model parts
3. Even early detection of implementation challenges or misunderstandings might turn out to be possible

We could expect implementations in varying sectors, for example in the IoT world where interoperability is achieved through the common SAREF ontology, maintained by ETSI. Another example is the standardization of procurement documents like the EN 16931 invoicing model.

6 Outlook and Future Work

These future work sections are recommendations for the future. Some could be done in IDSA as they are in the scope of IDSA's work. Other findings should be continued in other organizations in the wider community.

6.1 Topics Within IDSA Scope

Overarching model

IDSA working groups will continue to define the overarching model, including the core concepts for observability in Data spaces, and closely related topics to provide an overview and common understanding, such as how Data Planes fit into the observability part, and the relation to telemetry. This overarching model shall include the relationship to the lifecycle of data-sharing contracts.

Core data space observability operations

Examples such as data exchange/transfer metrics, contracting operations, identity and access management events, policy enforcement actions, catalog and metadata operations, consent management operations are considered as future work and may be explored in the next steps.

Meta observability agents for composite observability

Service Decomposition is about breaking down observability into specialized micro-services/Agents such as contract verification services, usage pattern monitoring, provenance tracking, compliance attestation, audit logging, and alert generation.

This topic is not within the core scope of the IDSA working groups. However, it may still be of interest to support an Impulse paper by members to explore business models that may be developed by service providers in data spaces based on observability.

6.2 Topics for the Wider Community

Migration paths

Transition from Clearing House to Observers is required, but not in scope of this document. This will be approached through individual dialogues with data spaces in need of this migration guidance, for the sake of keeping this document concise.

Implementation aspects

Further technical details and implementation guidance are not in the scope of this document. For instance, the definition of attributes and messages, where appropriate (i.e., open questions such as where is the use of such attributes agreed? Is it a matter of standardizing the protocols or part of an agreement negotiation process?) and implementation-specific observability such as Infrastructure health metrics, security monitoring, application performance, user experience metrics are best explored by parties working on specific implementations.

6.3 Research Areas

Research on „Confidentiality“ and „Relationship to compliance and automated compliance“ may be of interest to research organizations.

Confidentiality

Confidentiality controls on observability data and authority to access or use them need to be explored. Bilateral contracts on observability data as proposed earlier do not solve the problem of inference or collective harvesting of information that reveals secrets when a party is involved in multiple or all observability contracts in a data space. Such protection is necessary and beyond the scope or ability of any individual observability contract.

"Meta-observability" or what we might call "self-compliant observability" could be explored. Implementing a Self-Compliant Observability Framework as a recursive trust model where observers themselves are subject to the same or higher standards of transparency and compliance verification could be an interesting approach to explore by research organizations in next steps.

Relationship to compliance and automated compliance

Automated compliance refers to the use of technology, such as artificial intelligence (AI) and machine learning, to streamline and automate the process of adhering to regulatory standards, but also to data space rulebooks or data-sharing contracts. This involves continuously monitoring systems for compliance, replacing manual processes, and centralizing compliance tracking.

Compliance itself is the act of ensuring that an organization meets all relevant laws, regulations, standards, data space rulebooks, and contracts. Automated compliance enhances traditional compliance by minimizing human errors, saving time, and reducing risks associated with manual compliance management. It allows for real-time monitoring and reporting, making it easier for organizations to stay up-to-date with evolving regulations.

Observability in data spaces, as presented in this paper, can support automated compliance concerning the data-sharing contracts. The reconciliation of data-sharing contracts, or in general, policies and claims presented, which in general provide evidence of compliance, can be observed.

7 Conclusions

This document provides an initial overview of the concept of observability in data spaces. It outlines the motivation for introducing observability mechanisms, differentiates observability of data-sharing contracts from related concepts such as data provenance and traceability, and presents a classification of observability types and their implementation at both technical and business levels.

The paper identifies how observability can be integrated into the data-sharing lifecycle, during catalog, contract negotiation and transfer processes. It highlights potential roles for observers within data spaces and discusses the optional or mandatory use of observability based on governance requirements. The reference to semantic models and standardized schemas shows how observability can be implemented in a structured and interoperable manner.

Several use cases have been provided to illustrate practical applications of observability, including billing, regulatory and contractual compliance, AI training, and monitoring the adoption of standards. These examples demonstrate the relevance of observability in supporting trust, transparency, and accountability in data-sharing processes.

While the document focuses on key concepts and initial guidance, further work is required to define detailed specifications, develop implementation models, and address topics such as confidentiality, automated compliance, and future governance approaches. These aspects are outlined in the section on future work and may serve as a starting point for continued discussion within the IDSA community and the wider ecosystem.

Annex A. Additional details on the exemplary process

A.1. Activity Diagram for an exemplary process for setting up observability of a data-sharing contract

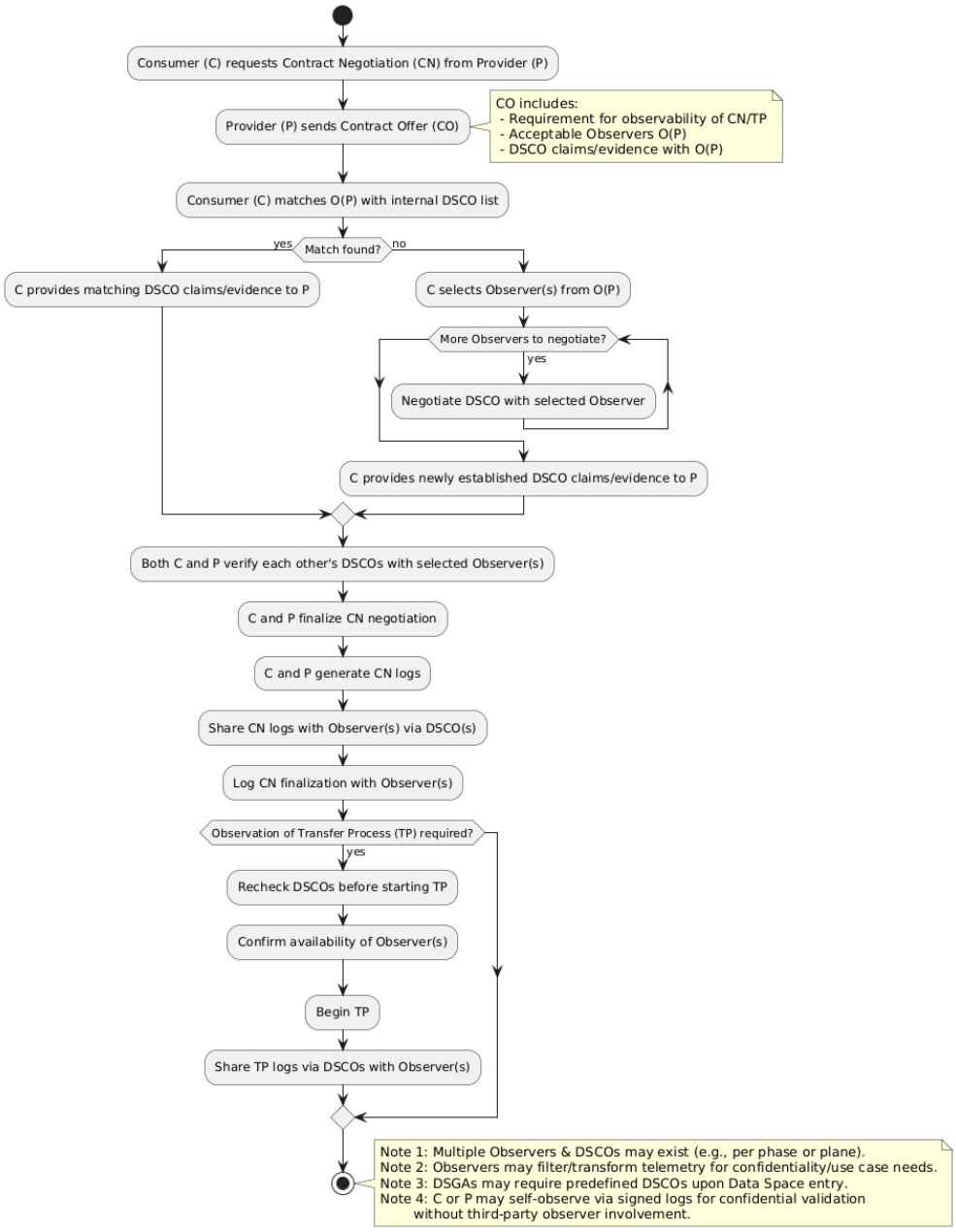

Figure 5 Activity Diagram for an exemplary process for setting up Observability of a data sharing contract

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