

EGI
2024

EGI
2024 EGI
EGI 202
EGI 2024
2024

Deep-Dive into Data Spaces: Choosing the Right Connector. Experiences from project TANGO, CEDAR, DIVINE and DATAMITE

Welcome and introduction

Mpampis Chatzimallis | The Lisbon Council

Giulia Giussani | IDSA

Today's agenda

Purpose of the workshop: shed light onto the projects and on the journey that led to the selection of the data space connector in the project

Time	Session	Speakers
15:15 – 15:30	Welcome and Introduction	Mpampis Chatzimallis, The Lisbon Council, TANGO
15:30 – 15:45	Introduction on data spaces and connectors	Giulia Giussani, IDSA
15:45 – 16:10	Project Pitches <ul style="list-style-type: none">• TANGO: Citizen-centric, privacy-preserving data sharing with energy-efficient operations - TANGO's white paper on connectors and future collaboration• CEDAR: Transparency and anti-corruption measures through better data governance• DIVINE: Open-source, resilient data spaces ensuring sovereignty and trust• DATAMITE: modular, open-source framework to enhance data monetization, governance, and interoperability	<ul style="list-style-type: none">• Moderator: Mpampis Chatzimallis, The Lisbon Council• TANGO: Renato Santana, EGI• CEDAR: Francesco Osimanti, The Lisbon Council• DIVINE: Sergio Comella, Engineering• DATAMITE: Alberto Berreteaga, Tecnalia
16:10 – 17:10	Deep-Dive into data space connectors	<ul style="list-style-type: none">• Moderator: Giulia Giussani, IDSA• TANGO: Kaitai Liang, TU Delft• CEDAR: Silvio Sorace, Engineering• DIVINE: Sergio Comella, Engineering• DATAMITE: Alberto Berreteaga, Tecnalia
17:10 – 17:45	Cross-project panel discussion and Q&A	Moderator: Giulia Giussani, IDSA Panelists: all speakers on-site & remotely connected
17:45 – 18:00	Key takeaways and closing	Giulia Giussani, IDSA

Introduction on data spaces and connectors

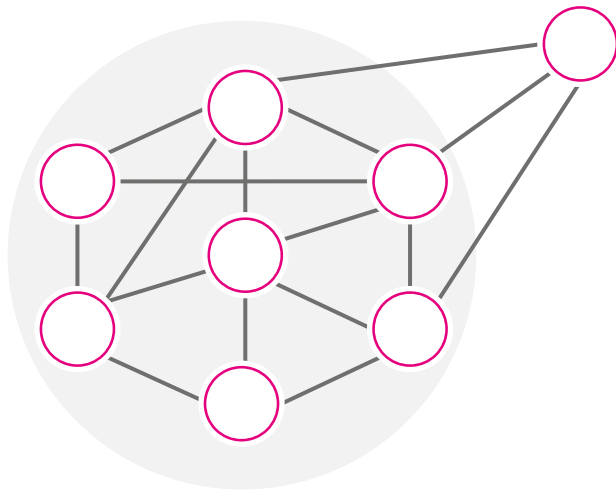
Giulia Giussani | IDSA

The data space approach

...connecting all kinds of data endpoints



A decentralized and dynamic data ecosystem:
with many-to-many interactions



A **data space** is the sum of all end points that are able to share data with each other.



- **Federated data architecture**: no physical data integration, leave data where it is
- **Interoperability**: no silos, no vendor-dependency
- **Data Sovereignty** and **traceability**
- **Trusted** participants



Plug in – or not

*Bringing interoperability and trust
to data ecosystems:*

*Participants in a data space need
to take an informed decision about
the trustworthiness of partners
and components used*



International Data Spaces Association

A not-for-profit organization creating a global standard for data sovereignty

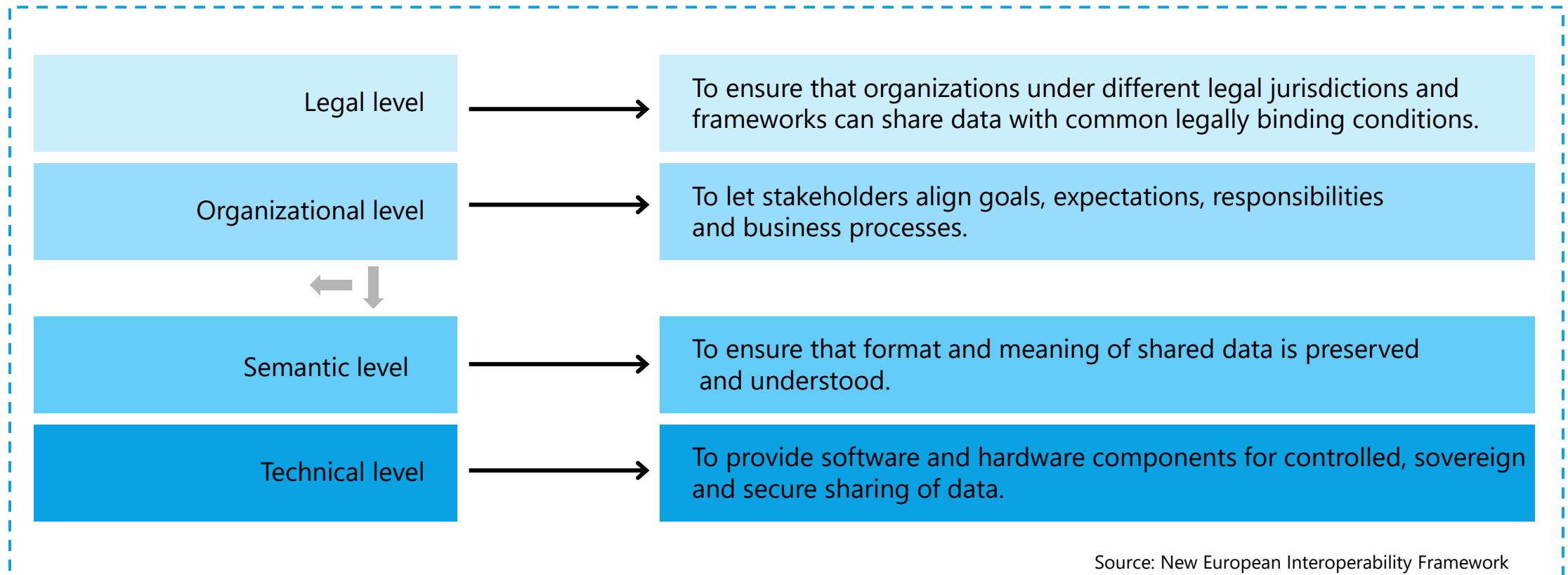
INTERNATIONAL DATA SPACES ASSOCIATION



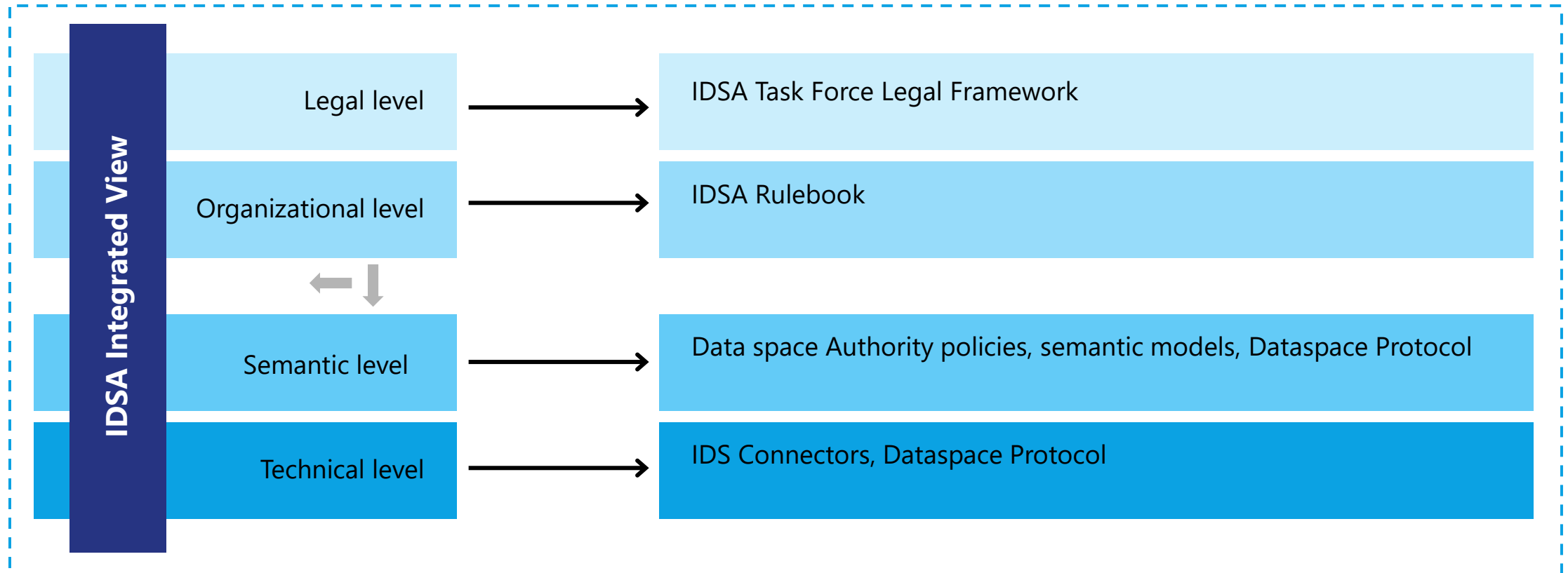
31 Countries

178

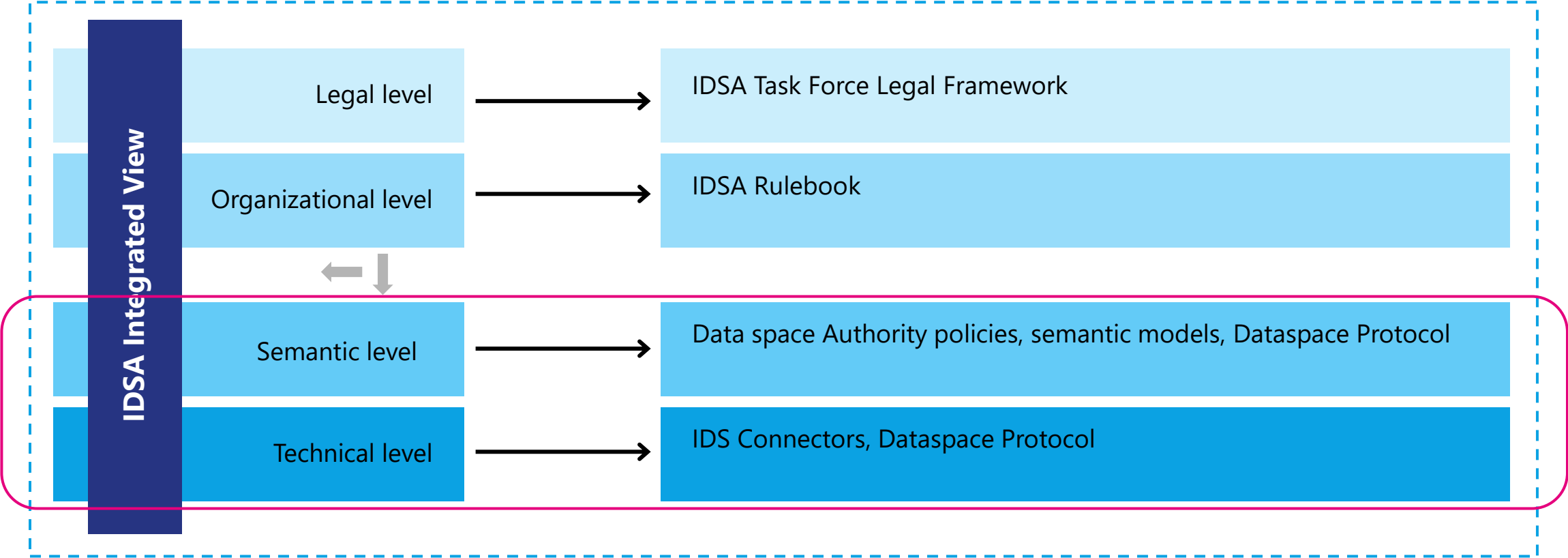
Layered model for interoperability

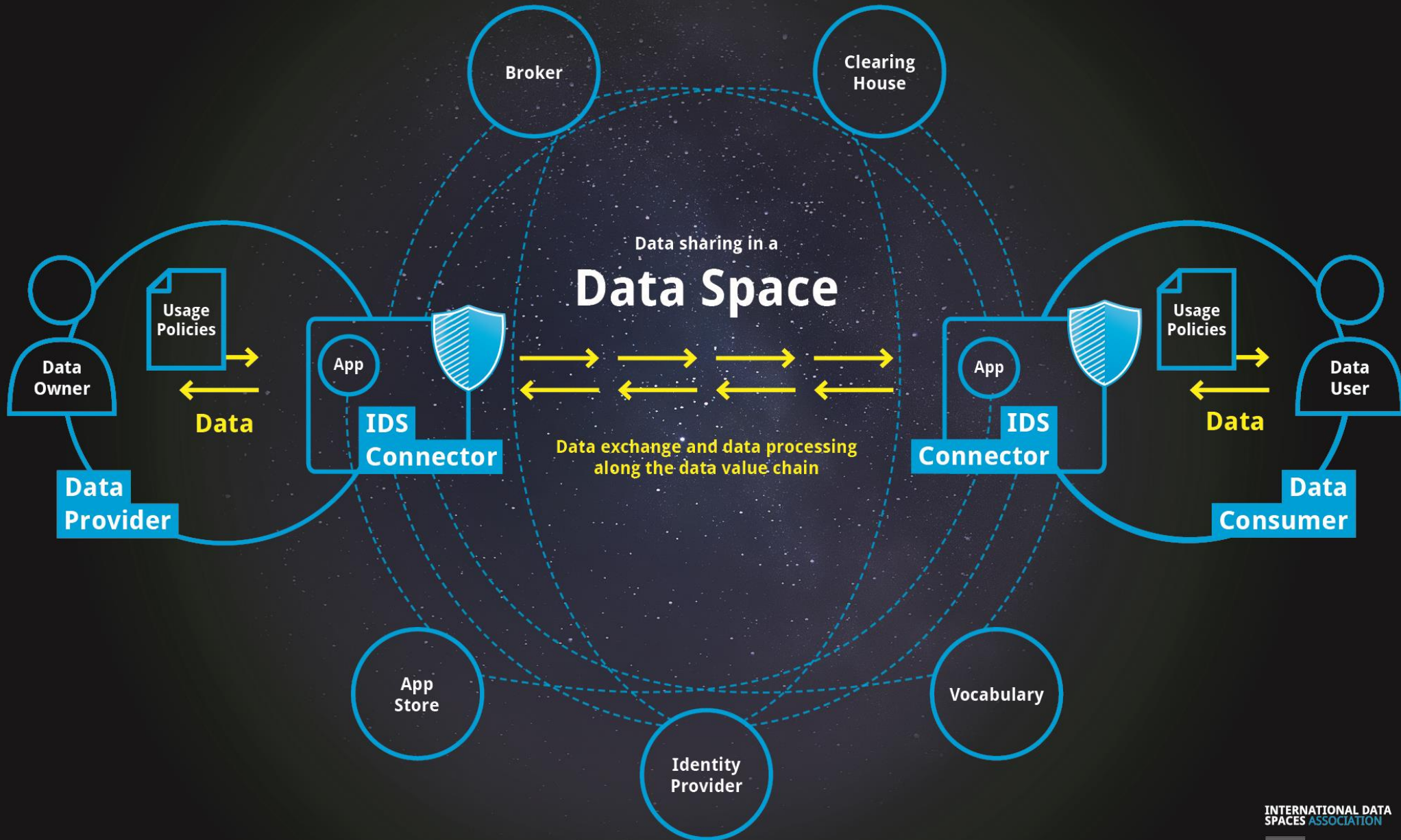


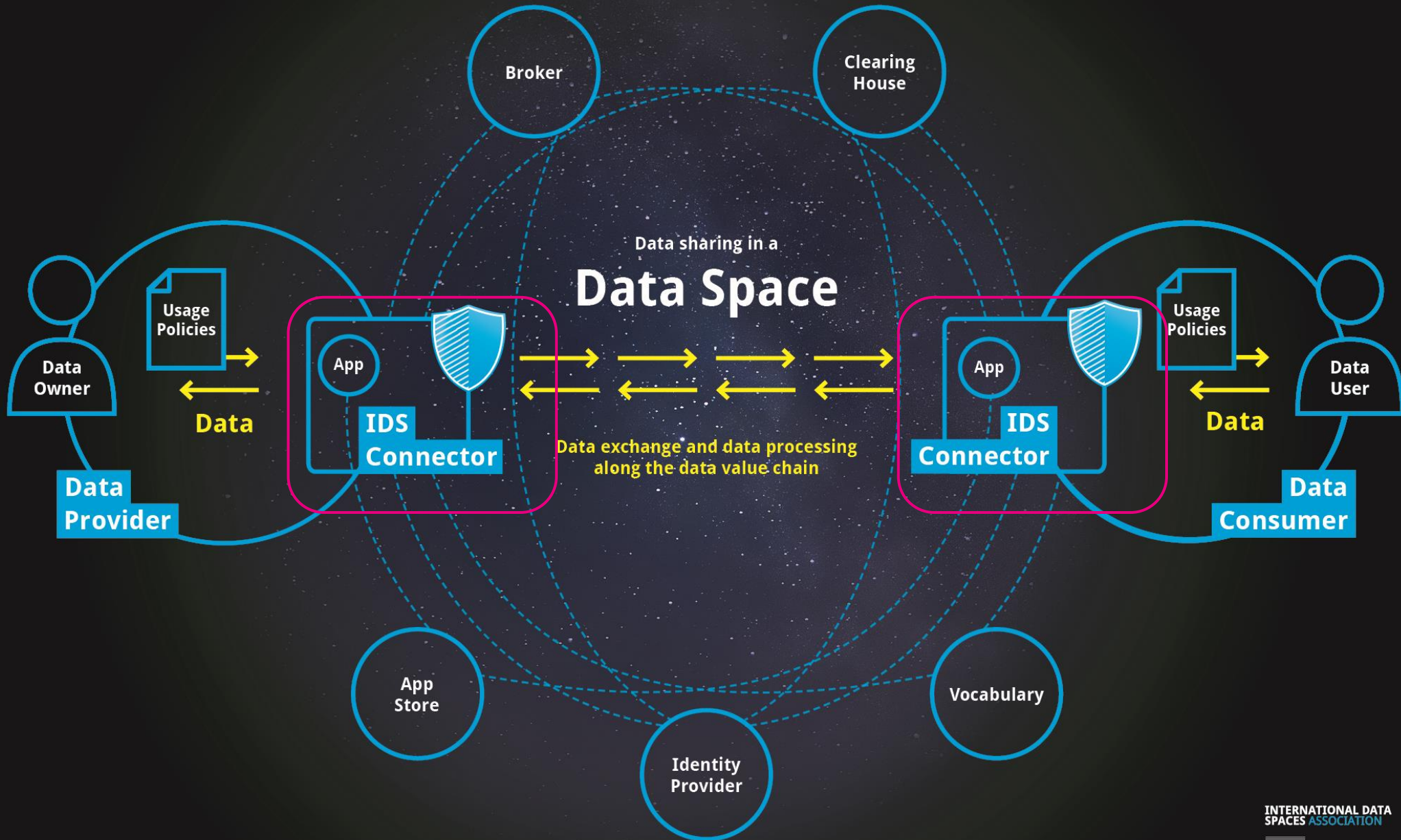
Layered model for interoperability



Layered model for interoperability





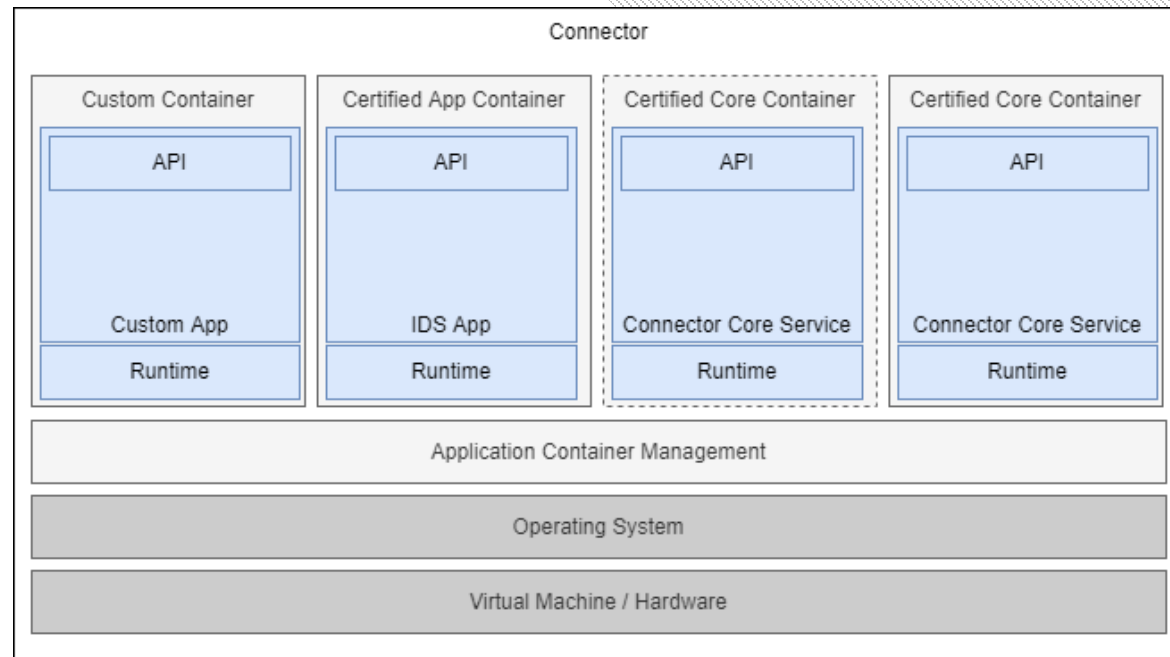


It all starts with the IDS Connector

The IDS Connector lays the basis for trustful data sharing



- » Each IDS Connector allows the exchange of data via the Data Endpoints it exposes.
- » IDS Connectors can be operated on-premises or in a cloud environment
- » The Connector consists of one or more computers/virtual machines, operating systems running on them, an Application Container Management, and the Connector Core Service(s) built on top of it.



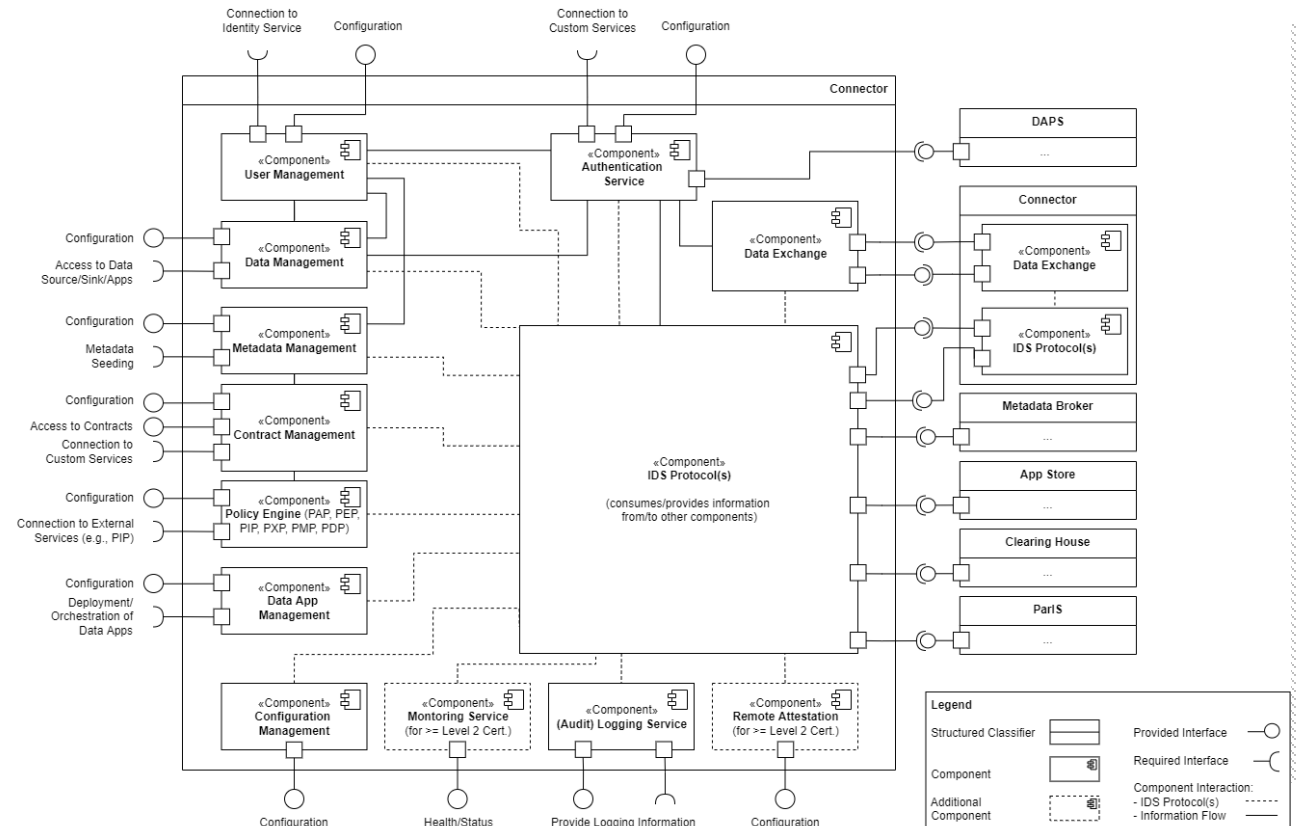
<https://docs.internationaldataspaces.org/ids-knowledgebase/ids-ram-4/layers-of-the-reference-architecture-model/3-layers-of-the-reference-architecture-model/3.5.0-system-layer/3.5.2-ids-connector#ids-connector-functionalities>

It all starts with the IDS Connector

The IDS Connector lays the basis for trustful data sharing



- » Some essential functionalities are needed in the Connector Core Service(s).
- » Examples: authentication service, remote attestation, data exchange,...
- » This UML deployment diagram from the IDSA RAM depicts each functionality as one component



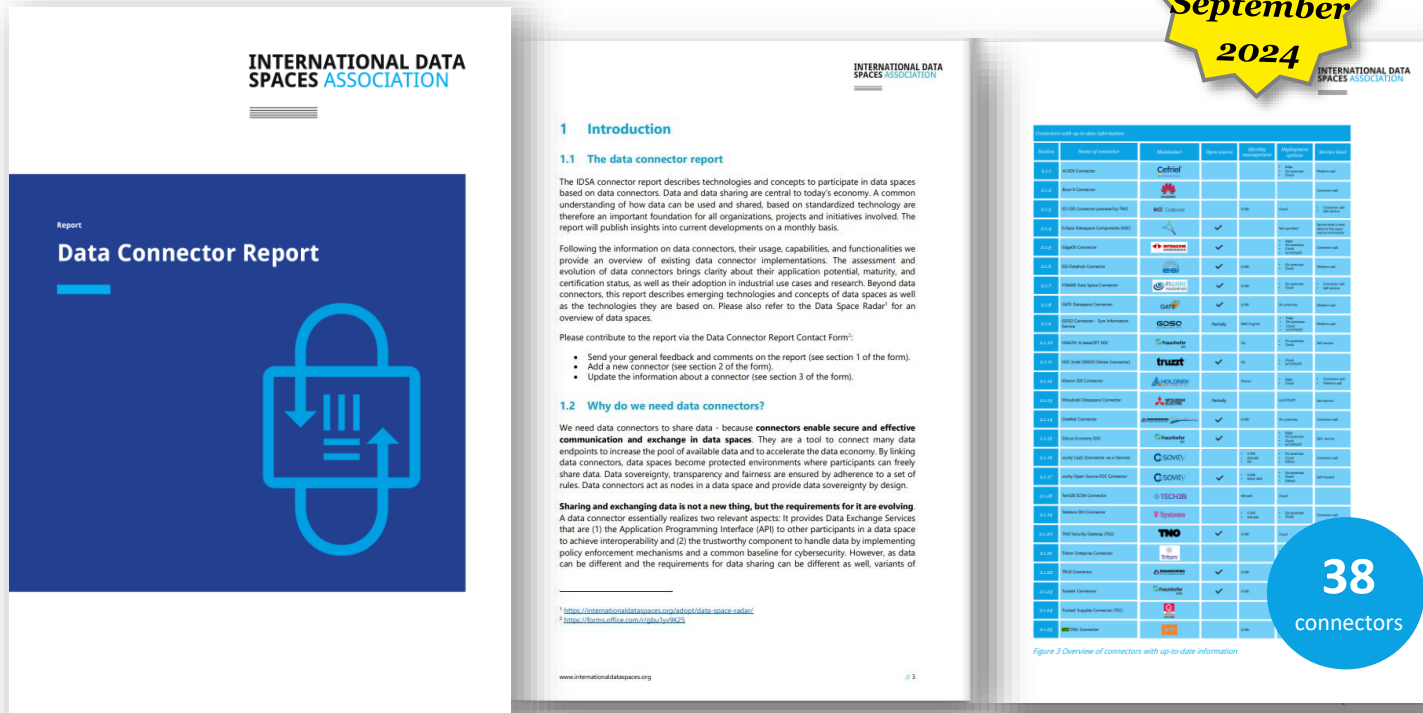
[https://docs.internationaldataspaces.org/ids-knowledgebase/ids-ram-4/layers-of-the-reference-architecture-model/3.5.0.system.layer/3.5.2.ids.connector#ids-connector-architecture](https://docs.internationaldataspaces.org/ids-knowledgebase/ids-ram-4/layers-of-the-reference-architecture-model/3-layers-of-the-reference-architecture-model/3.5.0.system.layer/3.5.2.ids.connector#ids-connector-architecture)

IDSAs Data Connector Report

Last edition of the Data Connector Report



Issue #16
September
2024



Why a Data Connector Report?

- To explain what data connectors are and why they are crucial in data spaces
- To provide transparency about the number of connector implementations available, their maturity and features, following their evolution over time
- To explain how data connectors can be technically interoperable
- To provide additional insights on related technologies and initiatives

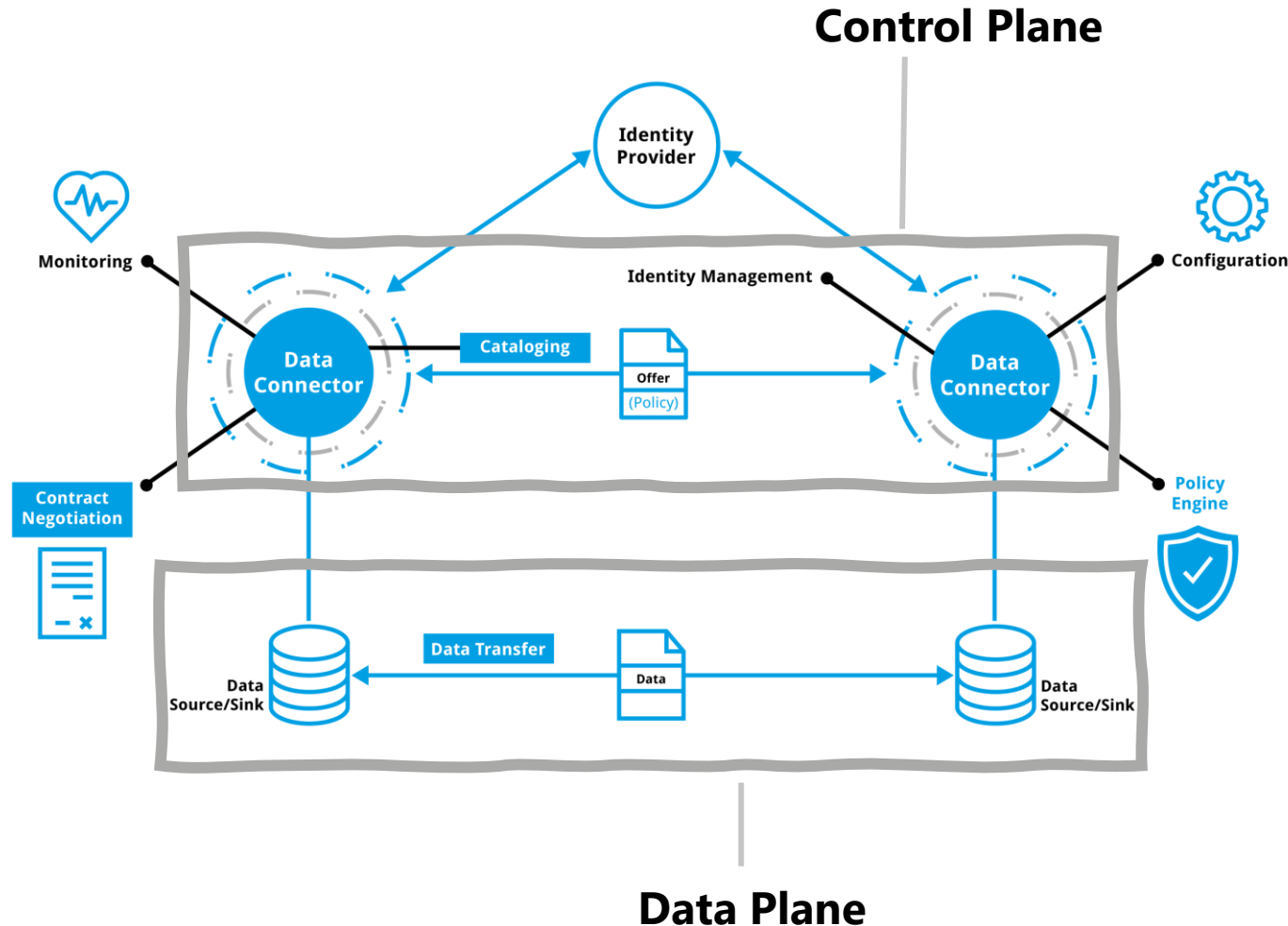
[Click here to take a look inside](#)



How to ensure interoperability?

The essence for interoperability

The Dataspace Protocol



Promotes seamless technical interoperability, while addressing certain aspects of semantic interoperability.



Enables standardized data exchange across different data space instances.



Provides flexibility and scalability through the separation of control plane and data plane.

Dataspace Protocol

Protocol's Structure



Catalog Protocol

- » Defines how data is listed and organized by the provider.
- » Makes data easy to find and understandable for potential consumers.
- » Ensures data is described in a consistent, standard format.
- » **You prepare and offer what is available**

Contract Negotiation Protocol

- » Facilitates the agreement on data usage terms between provider and consumer.
- » Defines how long, for what purpose, and under what conditions data can be used.
- » Provides a clear process to negotiate and finalize these terms.
- » **You negotiate and agree on how the data will be used**

Transfer Process Protocol

- » Manages the actual transfer of data once terms are agreed upon.
- » Ensures data is shared securely and follows the negotiated rules.
- » Supports different types of data transfers (e.g., one-time or continuous).
- » **You execute the data transfer according to the agreed terms**

Making the Dataspace Protocol an international standard

Making the Dataspace Protocol an international standard

Version 1.0 | June 2024

International standardization is fundamental to data space development

Harnessing the power of data is no longer optional but essential for success in business, politics, and society. However, to truly unlock its potential, we must handle data with utmost care. When collecting, processing, and sharing data, we need to prioritize trust and data sovereignty¹.

Data spaces are digital environments designed for the trusted sharing and management of data among various participants. Data spaces can enable more efficient implementations of advanced services and solutions based on data. Data spaces guarantee data sovereignty to participating individuals and organizations, allowing data holders to control the terms and conditions by which the data is re-used.

Data spaces have the potential to make business, economies, and life dramatically better. They are a paradigm shift for the way we share data – and they are a prerequisite to make the data economy and game changers like artificial intelligence happen.

But we at IDSA already back in 2017 observed that all these amazing things are not happening, at least not at the speed and scale we want and hope for. What was missing? A technology that integrates key processes common to all data spaces into a robust framework, following best practices, regardless of region, sector, use case, business model, or applicable legal regulations. Those are processes like exchanging information about what kind of data is available, negotiating a contract, and proceeding to the actual data exchange. What we

¹ <https://internationaldataspaces.org/why/data-sovereignty>

were missing is a data space protocol, better, the IDSA Dataspace Protocol². This challenge affected all organizations participating and providing services in data spaces.

Beyond common practices, we recognized that only standardized technology could provide organizations with sufficient confidence to invest in learning and implementing these new solutions. International standards are instrumental in addressing such challenges.

We benefit from standardization in other technological fields every day, whether we are aware of it or not. For example, when we fly to a foreign country and turn our mobile phone back on, we expect it to connect to the local communications network and make calls seamlessly. That is not magic, but the result of years of effort of academia, government, and industry converging to create and use standards.

The Dataspace Protocol – a baseline for data space interoperability

The Dataspace Protocol is not just a technical innovation; it is a strategic enabler for the data-driven economy. The DSP orchestrates the necessary steps for two or more parties to share data, including requesting a catalogue, negotiating a contract, and managing the transfer process. Its impact could be as profound as the protocols that shaped the Internet like the Internet Protocol (IP)³, transforming fragmented and isolated networks into a global, interconnected one. Similarly, the GSM standard⁴ revolutionized global mobile telephony. Just as these protocols laid the foundation for the Internet and modern communications, the Dataspace Protocol will lay the foundation for trustworthy and sovereign data sharing – unlocking new opportunities for collaboration and innovation in various sectors.



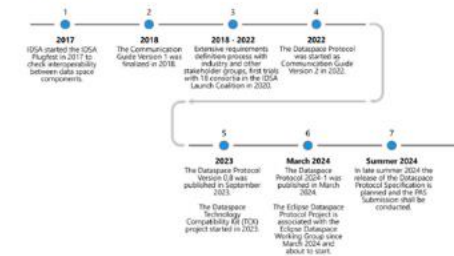
Dataspace Protocol – technical innovation and strategic enabler

² <https://internationaldataspaces.org/dsp/dataspace-protocol-overview>
³ <https://www.ietf.org>
⁴ <https://www.gsm-a.com>

Getting technology standardized is not straightforward; it requires choices. This paper describes the choices IDSA made and why they – in our view – form the best path to standardized data space technology. Throughout this journey, we address various challenges and advocate for clarity and transparency in developing this crucial technology.

Responding to the need for a specification for interoperable data spaces

The members of IDSA began designing technology for data spaces in our working groups and committees in 2017, so that the Dataspace Protocol could be ready to start the standardization process in 2024. What happened during those seven years? The diagram below offers a brief history of how we got here. The Dataspace Protocol in itself is quite lightweight – but it is the essence of the last seven years of IDSA work and therefore the IDSA Reference Architecture Model (IDS-RAM)⁵ and the IDSA Rulebook⁶ are at the core of it. We condensed all the knowledge, requirements and findings in this streamlined protocol so that it guarantees the benefit of much more comprehensive features that come with the whole IDSA framework.



Timeline of IDSA assets and activities leading to the Dataspace Protocol Specification

The European Union's Data Act⁷ will enter into force in September 2025 – just over one year from the publication of this paper. The new legislation specifies criteria for participants in data spaces to allow data to flow within and between data spaces, implying the use of harmonized interoperability standards. As it happened with other European legislation

⁵ <https://internationaldataspaces.org/ids/reference-architecture/>
⁶ <https://idsa.internationaldataspaces.org/ids/framework/idsa-rulebook>
⁷ <https://digital-strategy.ec.europa.eu/en/policies/data-act>



Project pitches

TANGO | Renato Santana, EGI

CEDAR | Francesco Osimanti, The Lisbon Council

DIVINE | Sergio Comella, Engineering

DATAMITE | Alberto Berreteaga, Tecnalia





Full name of the project

Digital Technology for Secure and Trustworthy Data Flows

Duration and budget

36 Months (1 September 2022 – 31 August 2025)

Budget: € 10 444 121,00

Goals

What is the aim of the project? Which are the outcomes? Is a specific domain in focus?

TANGO will establish a stronger cross-sector data sharing, in a citizen-centric, secure and trustworthy manner, by developing innovative solutions, for environmentally sustainable data management.

TANGO will evaluate the outcomes of the project with pilot demonstrations over several sectors such as: Smart Hospitality, Autonomous Vehicles, Smart Manufacturing, Public Administration, Retail Sector and Banking.

Exploitation

What are the benefits for the industry?

The TANGO platform provides industries with enhanced data security, privacy, and regulatory compliance, enabling secure and trustworthy data sharing. It promotes environmentally sustainable data management, helping businesses meet sustainability goals, while with decentralized, user-friendly solutions, TANGO reduces operational costs and improves efficiency.

When will the results be available?

TANGO is entering the evaluation phase and is currently deploying its solutions to pilot sites, with results being expected in a year from now.

Project website: <https://tango-project.eu/>

Contact details: renato.santana@egi.eu



Full name of the project

Common European Data Spaces and Robust AI for Transparent Public Governance

Goals

What is the aim of the project? Which are the outcomes? Is a specific domain in focus?

CEDs and EU Data Strategy

Framework/Architecture to build upon CEDs and exploit their Data Sharing potential to support the uptake of a better, more accountable and transparent Public Governance in Europe.

Duration and budget

36 months (1 January 2024 – 31 December 2026)

Budget: € 8 999 550,00

Exploitation

What are the benefits for the industry?

The benefits/impact from CEDS will be huge for the industry and the overall society, since the value generated by the access to data will be creating new streams of business and new economic opportunities . Specifically from CEDAR, the transparency of public funds and the reduction of corruption channels will create fair entrepreneurial ecosystems that will permit companies to thrive.

When will the results be available?

The results will be available in few years from now, as the CEDS will be populated with data and the technologies for interconnectivity will be maturing.

Project website: <https://cedar-heu-project.eu/>

Contact details: francesco.osimanti@lisboncouncil.net



Full name of the project

Demonstrating Value Of Agri Data Sharing For Boosting Data Economy In Agriculture

Goals

What is the aim of the project? Which are the outcomes? Is a specific domain in focus?

The DIVINE project aims to enhance **data sharing** and **collaboration** in the agricultural sector by addressing challenges like **interoperability** and **data ownership**.

Its goal is to build a collaborative **agri-data ecosystem** that improves decision-making and demonstrates the value of shared data.

The project focuses on agriculture, driving innovations in **precision farming** and **sustainable practices** while contributing to **policy-making** and the adoption of **digital technologies**.

Duration and budget

36 Month (October 2022 - September 2025)

Budget: € 3 954 797,00

Exploitation

What are the benefits for the industry?

The benefits for the industry from the DIVINE project include enhanced **efficiency** and **competitiveness** through improved **data sharing** and **collaboration** within the agricultural sector. By leveraging data-driven decision-making, the project aims to create **new business models**, foster **innovation**, and promote **sustainable farming practices**, ultimately benefiting the entire agri-food value chain.

When will the results be available?

The results of the project are expected to be available by **September 2025**, as the project runs from **October 2022 to September 2025**.

Project website: <https://divine-project.eu/>

Contact details: sergio.comella@eng.it



Full name of the project

DATA Monetization, Interoperability, Trading & Exchange

Goals

DATAMITE delivers a modular, open-source and multi-domain Framework to improve DATA Monetizing, Interoperability, Trading and Exchange, in the form of software modules, training and business materials for European companies, empowering them to become new relevant players in the data economy.

DATAMITE 3 Use Cases (UCs) and 6 pilots differentiate two key parts, how organizations manage their data and how it is shared.

In all UCs, data governance, quality or security modules are used, aiming at improving how data is managed within companies, how good is data, or interoperability.

Pilots context: Large Corporations, Energy Data Spaces, eDWIN + MISTRAL

Duration and budget

01/01/2023 - 31/12/2025
EC Horizon Europe
11,200,000 €

Exploitation

Data monetization: Using your data to add to or increase your revenue stream

Internal data monetization is the method of using data and analytics to make informed business decisions that turn into measurable improvements

External data monetization is the method of creating a product or service using your internal data assets and selling them to a third party

Open-source framework - Collection of open-source modules for data quality, security, sharing and governance

State of the art dataspace technologies, connectors, portals, etc

Framework Architecture – Technical implementation of a tool for data monetisation

Business model strategies for data monetisation (e.g. for data scientists)

Open datasets (e.g. forecast and weather data)

Training material – How to use DATAMITE's products (upskill personnel)– How to utilise available datasets (upskill personnel)

Guidelines/Training towards non-monetary benefits

Data monetization strategies and maturity model

Legal framework

Engagement with communities sharing experiences and expertise from DATAMITE

Data Support Tools (Harmonisation, Anonymisation, etc)

Deep-dive on data space connectors

TANGO | Kaitai Liang, TU Delft

CEDAR | Silvio Sorace, Engineering

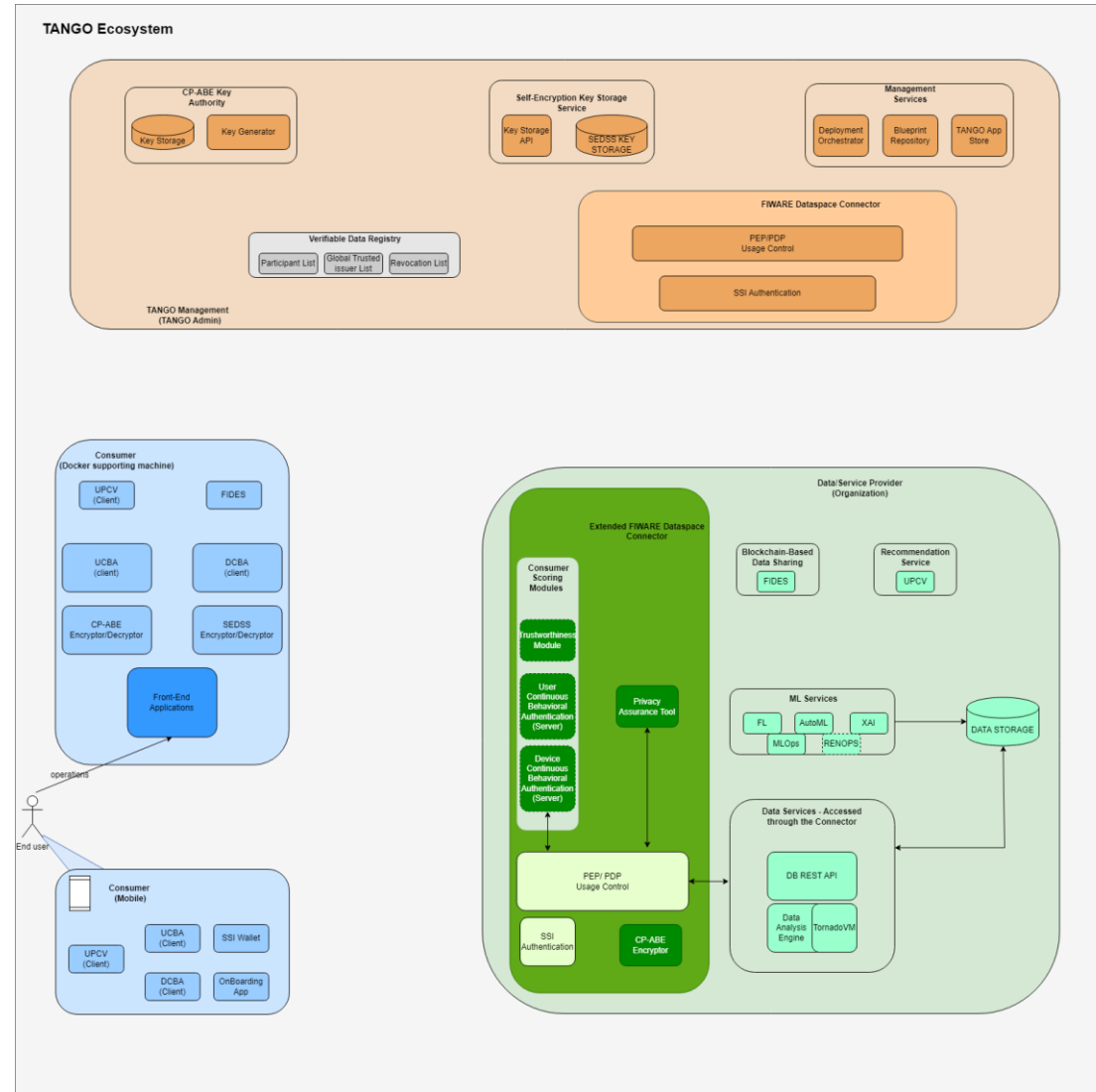
DIVINE | Sergio Comella, Engineering

DATAMITE | Alberto Berreteaga, Tecnalia





Architecture





Connector

Connector name

TANGO Connector

Details on the connector

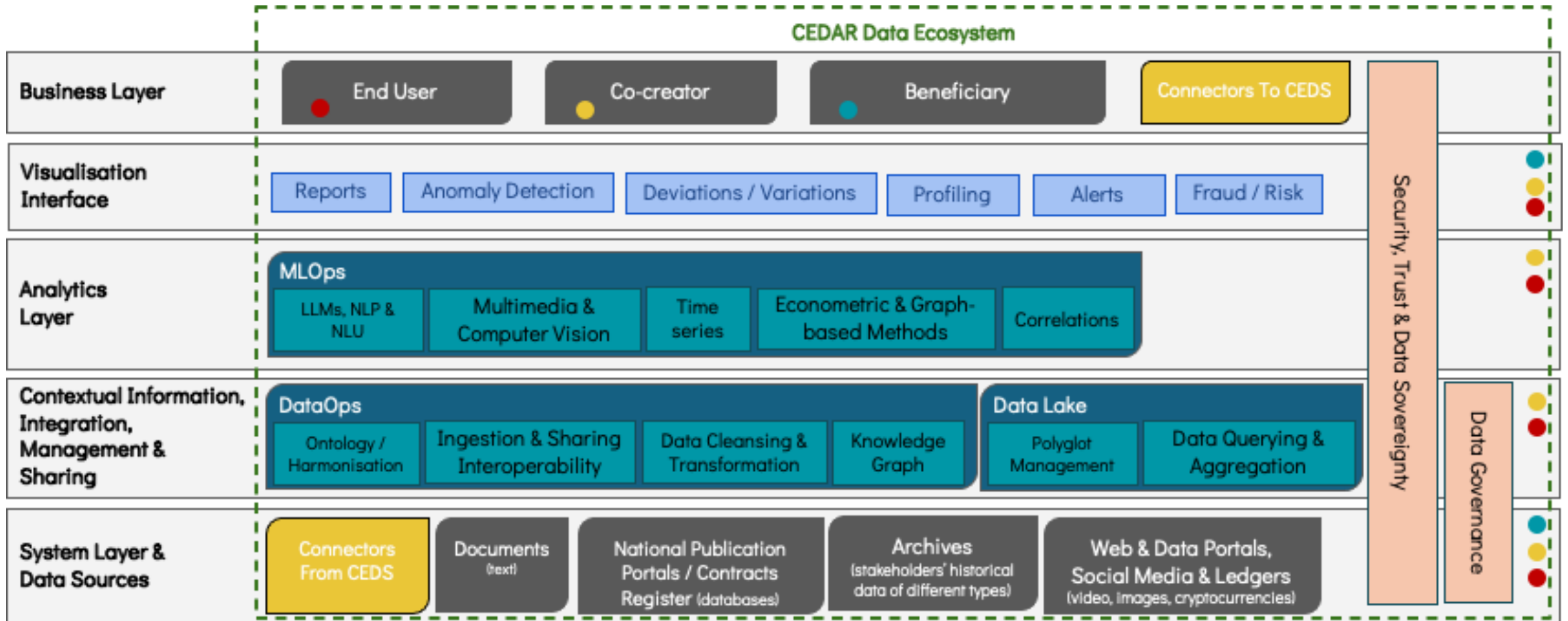
The TANGO Connector enhances the FIWARE Dataspace Connector by replacing the existing PDP/PEP modules with its own advanced modules, designed to evaluate the trustworthiness and behavior of requesters during the access request process. This evaluation is conducted using attribute-based policies and a range of trustworthiness scores. Additionally, TANGO integrates CP-ABE encryption, adding a robust encryption layer to secure the data exchanged through the connector.

Features at a glance

Feature	Details
Maturity	TRL 6
License	Partially open source. The FIWARE Dataspace Connector is open source. The TANGO developments are not
Identity management	<ul style="list-style-type: none">▪ Decentralized - did:web▪ Decentralized - SSI
Access control	<ul style="list-style-type: none">▪ OAuth
Usage control	<ul style="list-style-type: none">▪ Data Consumer▪ User Role
Communication protocol	<ul style="list-style-type: none">▪ HTTPS REST APIs
Graphical user interfaces	No



Architecture





Architecture

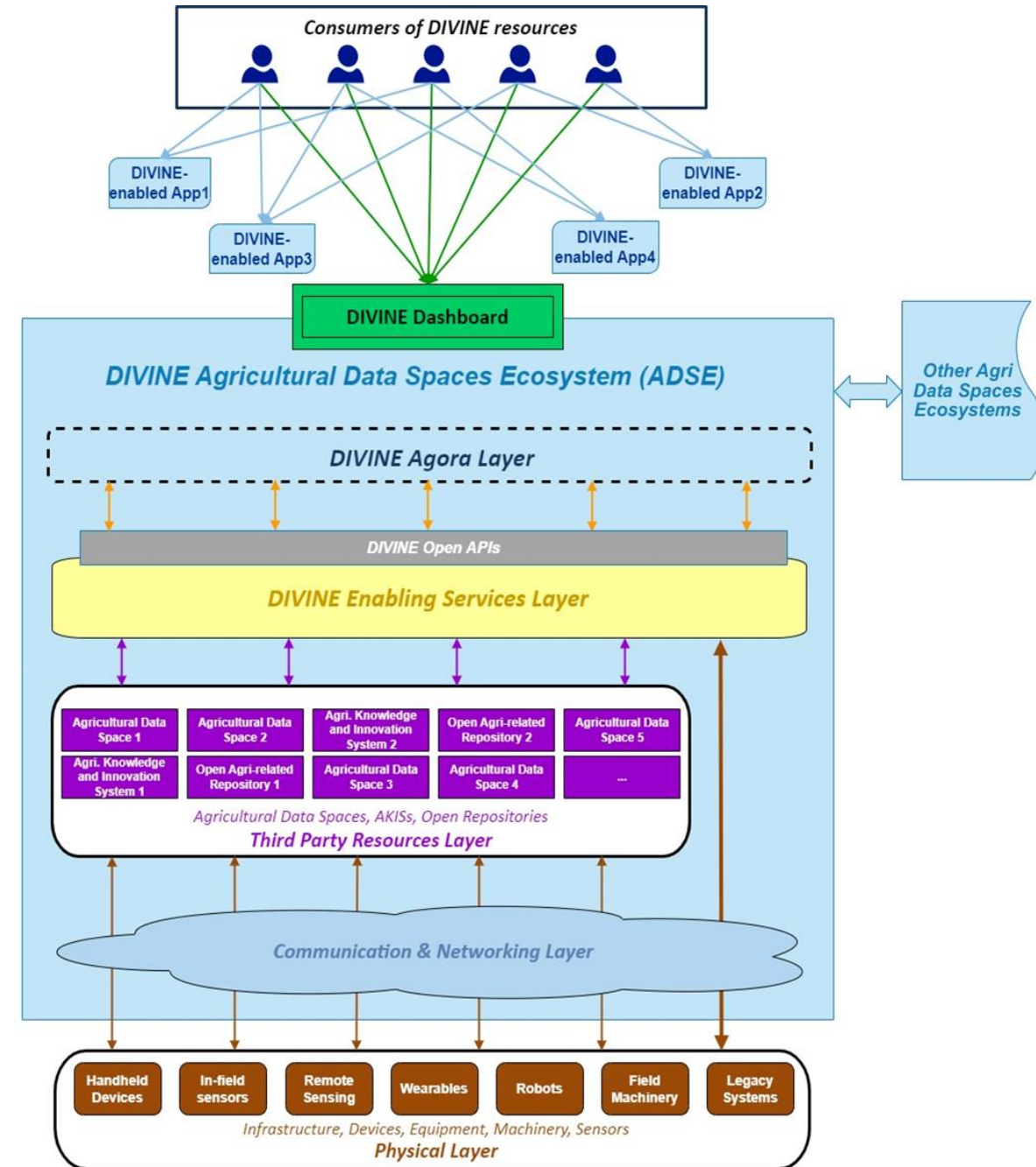
DIVINE AGORA LAYER

Developers and stakeholders can register their capabilities and resources as **DIVINE Providers**.

DIVINE Consumers can then browse the DIVINE Agora to find resources and capabilities that match their needs.

DIVINE Enabling Services Layer

It provides modular services for stakeholders





Connector

Connector name

Enhanced True Connector

Details on the connector

The connector used in DIVINE is based on the first version of the True Connector, modified to follow the Dataspace protocol.

Thanks to this, its unique value lies in the integration of agri-specific functionalities, providing seamless interoperability within the Agricultural Data Spaces. This improves data exchange between different stakeholders and provides a secure, transparent and efficient solution tailored to the agricultural sector.

Features at a glance

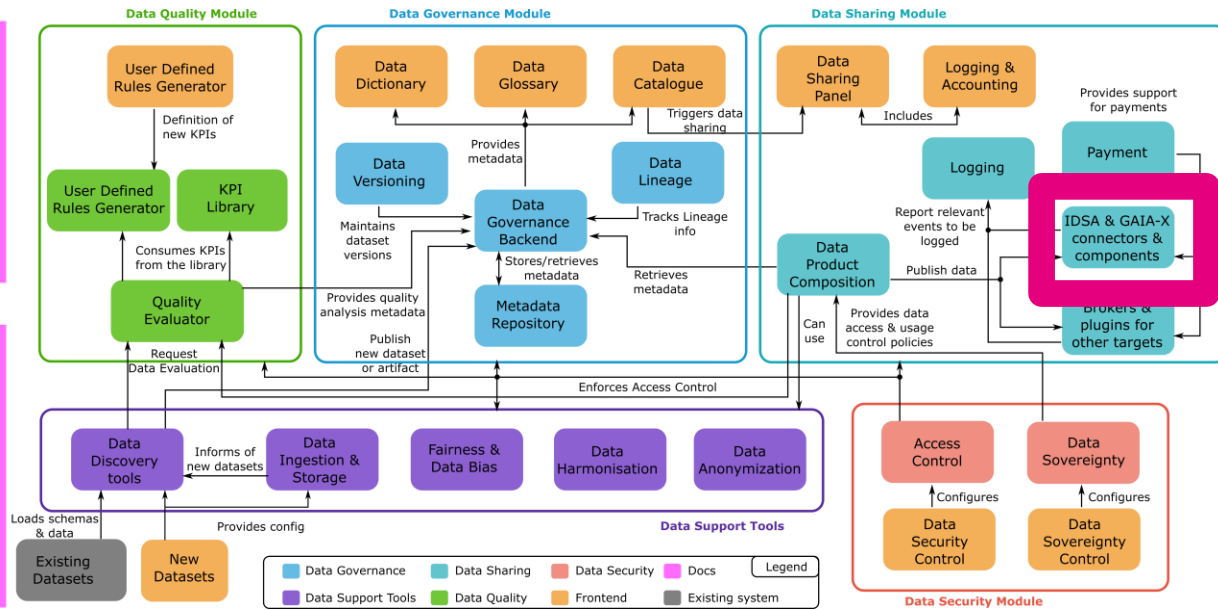
Feature	Details
Maturity	TRL 3
License	The AGPL-3.0 is an open-source license that requires users to provide access to the source code, including when the software is used over a network, ensuring user freedom in hosted environments.
Identity management	<ul style="list-style-type: none">Decentralized - did:webDecentralized - SSI
Access control	<ul style="list-style-type: none">OAuth
Usage control	<ul style="list-style-type: none">Platoon Usage Control
Communication protocol	<ul style="list-style-type: none">HTTPS REST APIsIDSCPv2SFTPOther implementable protocols
Graphical user interfaces	Yes



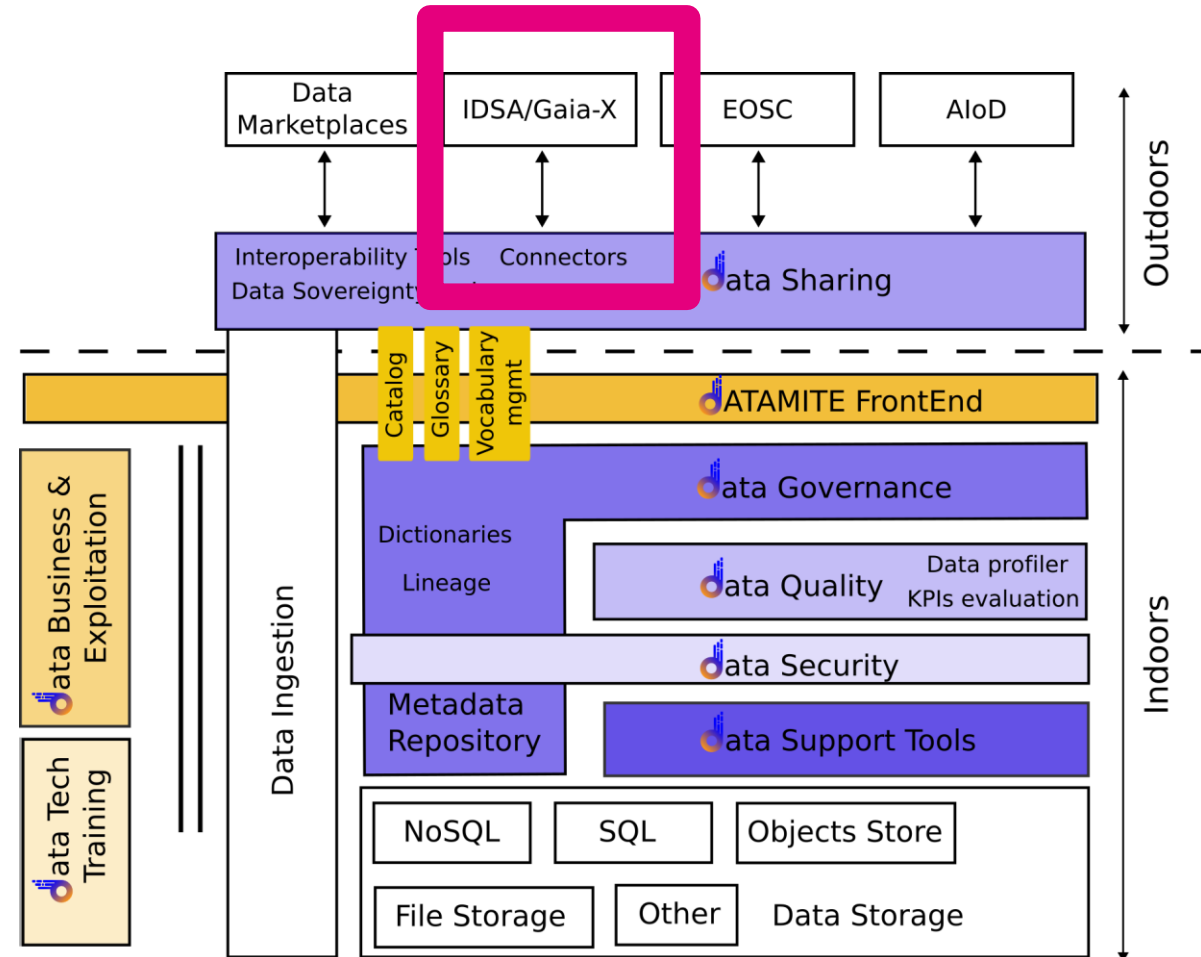
Architecture

data Business & Exploitation material

data Tech Training material



«thematic» architecture view



«planes» (indoors/outdoors) architecture view



Connector

Connector name

Eclipse Dataspace Components – Framework

Details on the connector

- https://internationaldataspaces.org/wp-content/uploads/dlm_uploads/IDSA-Data-Connector-Report-84-No-16-September-2024-1.pdf
- <https://github.com/eclipse-edc/Connector>
- <https://projects.eclipse.org/projects/technology.edc>
- <https://github.com/eclipse-edc/Samples/blob/main/basic/basic-01-basic-connector/README.md>
- EDC as open-source approach
- Well known, used, maintained, Gaia-X approach, catalogs and **extensions**

Features at a glance

Feature	Details
Maturity	TRL 8-9
License	Open source (Apache 2.0)
Identity management	No provided
Access control	Subject to data planes
Usage control	Subject to data planes
Communication protocol	Dataspace protocol 2024-01 (HTTPS) Transfer protocol: Out-of-band utilizing data planes without determined protocol bindings (data planes to be added and not part of EDC)
Graphical user interfaces	No

Cross-project panel discussion and Q&A

Moderated by Giulia Giussani | IDSA



Other connectors evaluated

Connector name	Pros	Cons
Dataspace Connector	<ul style="list-style-type: none">• Alignment with initially defined architecture• Support for deploying apps (better integration with rest TANGO components)• Support for implementing data flows inside the connector (Apache Camel)	<ul style="list-style-type: none">• No Decentralized Identity – SSI• No longer maintained
EDC Connector	<ul style="list-style-type: none">• Decentralized Identity Management - Offers integration endpoints for SSI• Modular, highly extensible and customizable• Very active community and well-supported• Data plane transfers support multiple protocols such as HTTP, Kafka, etc.	<ul style="list-style-type: none">• Mostly a toolbox rather than an integrated framework• Usage policies must be implemented• Complex setup and configuration• Limited Documentation
TRUE Connector	<ul style="list-style-type: none">• Advanced usage control policies with MYDATA (roles, time interval, duration, number of usages, etc.)• Modular, allows for independent extensions in either control plane, data exchange or usage control components• End-to-end data traceability	<ul style="list-style-type: none">• Decentralized Identity Management is not supported



Other connectors evaluated

Connector name	Pros	Cons
EDC	<ul style="list-style-type: none"><li data-bbox="919 511 1327 546">• Early integration of DSP	<ul style="list-style-type: none"><li data-bbox="1653 494 2226 565">• It doesn't strictly adhere to the DSP specifications.



Other connectors evaluated

Connector name	Pros	Cons
Data Space Connector	<ul style="list-style-type: none"> • See below 	<ul style="list-style-type: none"> • See below
Other IDS Connectors	<ul style="list-style-type: none"> • Some mature connectors • IDS standards 	<ul style="list-style-type: none"> • Different open-source mature elements • Different evolution path / steps (Data Space Protocol implementations) • Usage • Extensions / plugins

Differences between DSC and the EDC

DSC

Synchronous Communication

Deprecated

Not Scalable

Ready-to-Use-IDS-Connector

Supports only IDS

Monolithic Software Package

No Extensions

EDC

Asynchronous Communication

Large open-source community

Scalable

Framework to build an own Connector

Support for IDS, as well as other standards

Modular Architecture

Support for Extensions

Key takeaways and closing

Giulia Giussani | IDSA

Key takeaways (1)



- IDS Connectors are crucial in data spaces and a variety of connectors are arising from the market
- Different projects need connectors with different functionalities, but a common foundation for interoperability in data spaces is greatly needed
- The essence of interoperability is the Dataspace Protocol, which is on its way to international standardization
- New European legislations: the Data Act mandates essential requirements for participants in data spaces to ensure that data can flow freely. This requires harmonized interoperability standards



Key takeaways (2)

- Follow the projects LinkedIn accounts to stay up-to-date on the developments
- Connect with the experts after the workshop, in person or via the email addresses
- Join the data space pioneers: support IDSA and become a Member



**INTERNATIONAL DATA
SPACES ASSOCIATION**





Giulia Giussani

Digital Innovation Manager | IDSA



Emil-Figge-Straße 80, 44227 Dortmund



giulia.giussani@internationaldataspaces.org



[Giulia Giussani](#)