Fraunhofer ISST

Prof. Dr. Boris Otto · 31 January 2022

Federated Data Spaces

public

Table of Contents

public

Catena-X is creating a data ecosystem for the automotive value chain

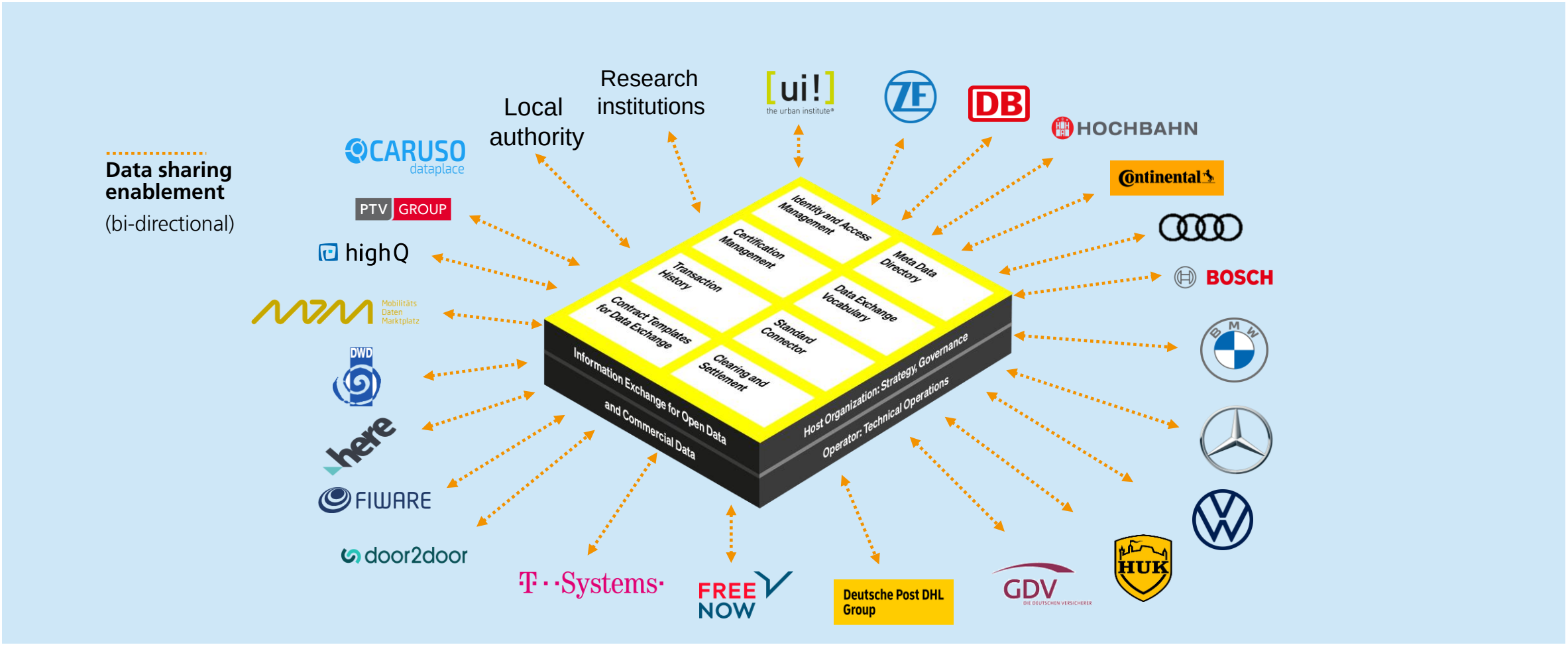
- Ecosystem use cases
  - Demand and capacity management
  - Circular economy
  - End-to-end compliance (Supply Chain Law etc.)
  - ...
- Design principles
  - Provide user-friendly environment for collaborative use of end-to-end data chains
  - Openness and non-discriminatory access for all market participants (free and open-source software)
  - Certified, network-based applications and solutions based on open standards
  - Support project and association structure

NB: Consortium refers to partners of the project funded by the German Ministry for Economic Affairs BMWi.

public

The Mobility Data Space facilitates a data sharing community for innovative mobility services

Mobility Data Space
Data Sharing Community



NB: The distributed ecosystem platform is an implementation of the IDS-RAM.

public

Fraunhofer ISST

Today, data ecosystems are emerging in literally all domains to collaboratively achieve customer innovation through data sharing

Finance & Insurance

Mobility

Healthcare

Industry 4.0

Energy

Aerospace

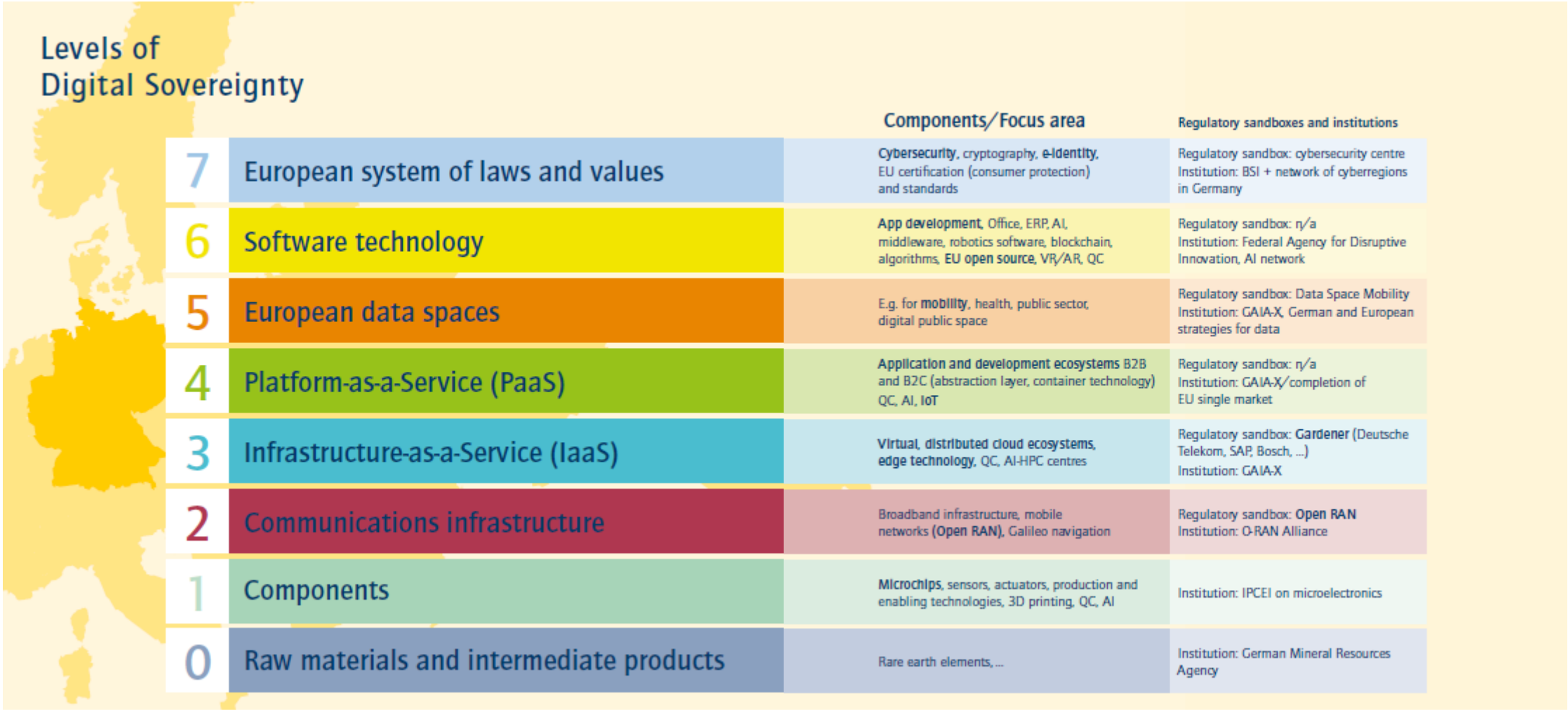
A multilateral form of organizing for joint customer innovation

Balance of the viability of the ecosystem as a whole and of its individual members

Image sources: Fotolia (2018); ubimet (2020); Viro (2020). **NB:** As Presented at GAIA-X Summit on 18 November 2020.

public

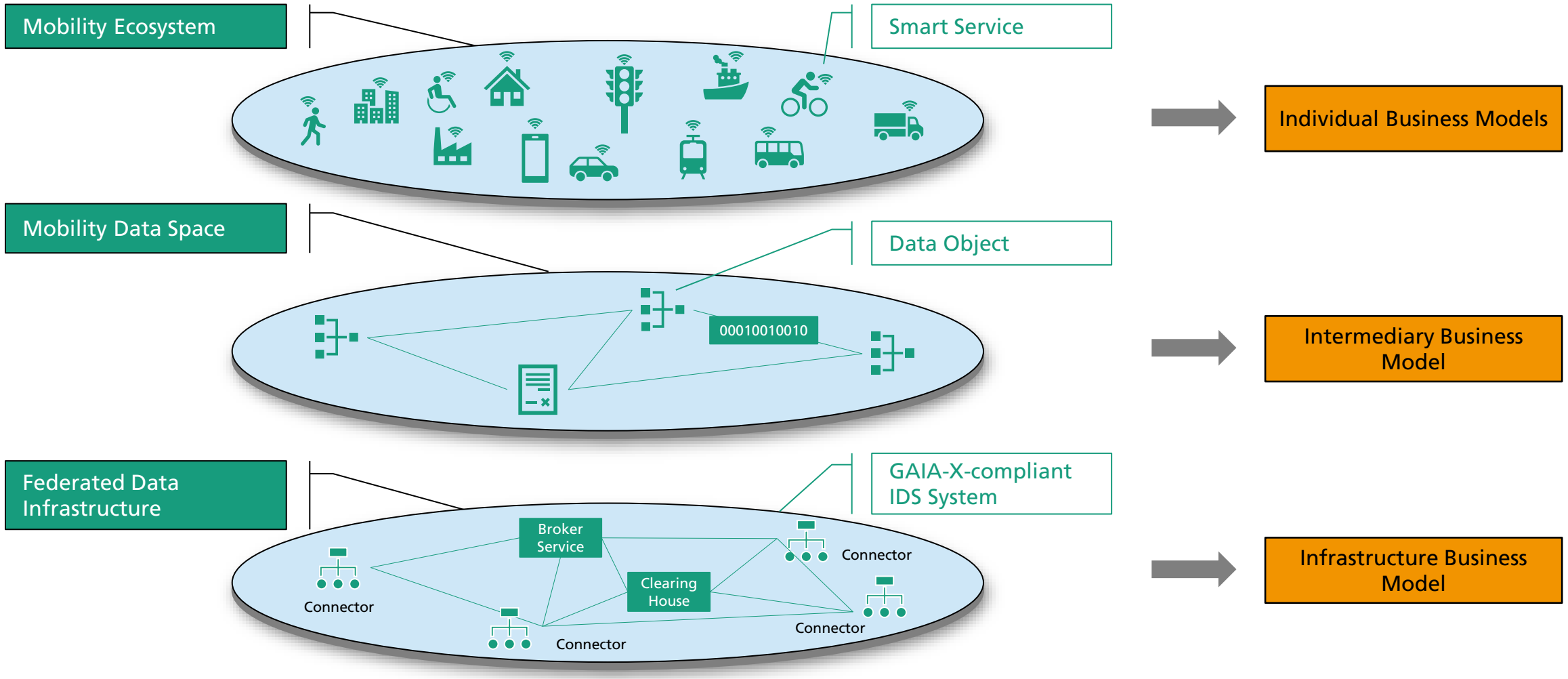
Data spaces are part of a larger architecture stack addressing digital sovereignty



Source: Kagermann et al. (2021).

public

Data spaces rest on a federated data infrastructure and enable data ecosystems



public

The notion of data spaces has evolved from the Linked Data research community

- General Design Principles[1]
  - No physical data integration, leave data where it is (→ **federated data architecture**)
  - No common schema required (→ integration foremost on **semantic level through vocabularies**)
  - Data networking, data visiting and data co-existence
  - Nesting and overlaps possible (→ **ecosystem of data spaces**)
- Additional IDS[2] Design Principles
  - Data sovereignty and traceability
  - Trusted participants

1) See (Franklin et al., 2005), (Halevy et al., 2006).
2) See (IDSA, 2019a)

public

Today, the understanding of the data space notion is at least threefold

Technology

Data Integration Concept

- No physical data integration
- No common schema required
- Data networking, data visiting and data co-existence
- Nesting and overlaps possible
- Data sovereignty and traceability
- Trusted participants

Business

Form of Collaboration on Data

- Business collaboration format
- Multilateral organization for a data sharing purpose
- Shared goals of participants regarding the exchange and sharing of data
- Decision making body for collective data governance

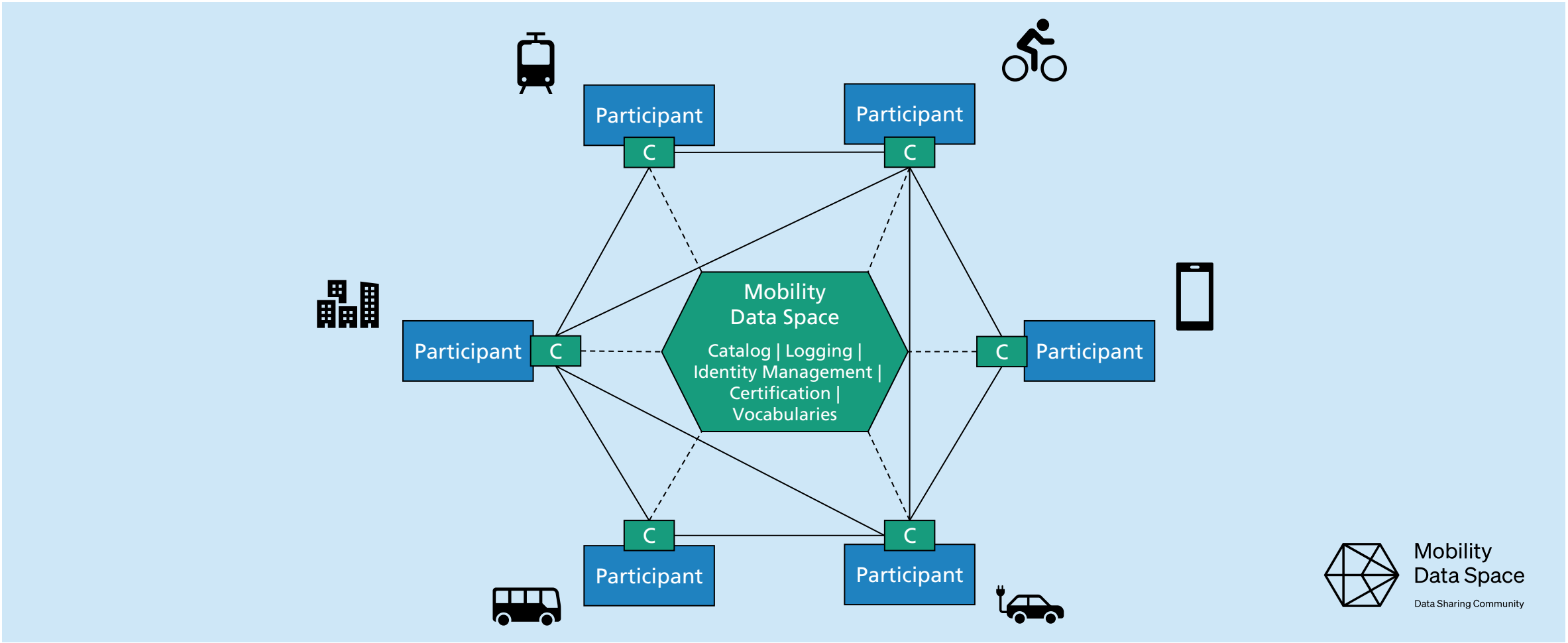
Legal

Object of Regulation

- Instrument for the implementation of the European and member states' data strategies
- Potential object of application of EU Data Governance Act
- Object of application of European values for data sharing and data sovereignty (cf. GAIA-X Policy Rules)

public

The Mobility Data Space is a good example for a mulit-stakeholder data ecosystem aiming to share various types of mobility data.



Legend: ---- Metadata; —— Data; C – Connector. Source: Otto (2022, soon to be published).

public

Business requirements for data spaces are relatively clear while economical and governance implications are still somewhat open

Business Requirements[1]

- **Ecosystem**
  - Ecosystem must comprise data providers, data consumers, and intermediaries
  - Ecosystem must be **open**, but participants and software endpoints must be **certified**
  - Ecosystem must allow different data to be treated differently, depending on its security classification and its nature as an economic good
- **Data**
  - Data **rights** must be clarified and protected
  - Data **heterogeneity** must be supported
  - Data flow **traceability** must be possible
  - Data **usage conditions** must be manageable and remotely enforceable
- **Platform**
  - Platform must allow integration and use of existing technologies and **standards**
  - **Trust** and **security** must always be ensured

Economical and Governance Implications

- **Financing and Funding**
  - Infrastructure-like funding?
  - Usage-based fees?
- **Data Space Ownership and Operations**
  - Public-private partnership?
  - Cooperative society?
  - Industry consortium?
- **Incentive Systems for Providers and Users**
  - »Tokenomics«?
  - Early adopter advantages?
- **Data Governance and Data Economy**
  - Policies as »Terms and Conditions for Data«?
  - Instrument to comply with EU Data Governance Act?
  - Technical means for data usage contracts?
  - Technical means to build federated data trusts?

1) Source: (Otto & Jarke, 2019).

public

Table of Contents

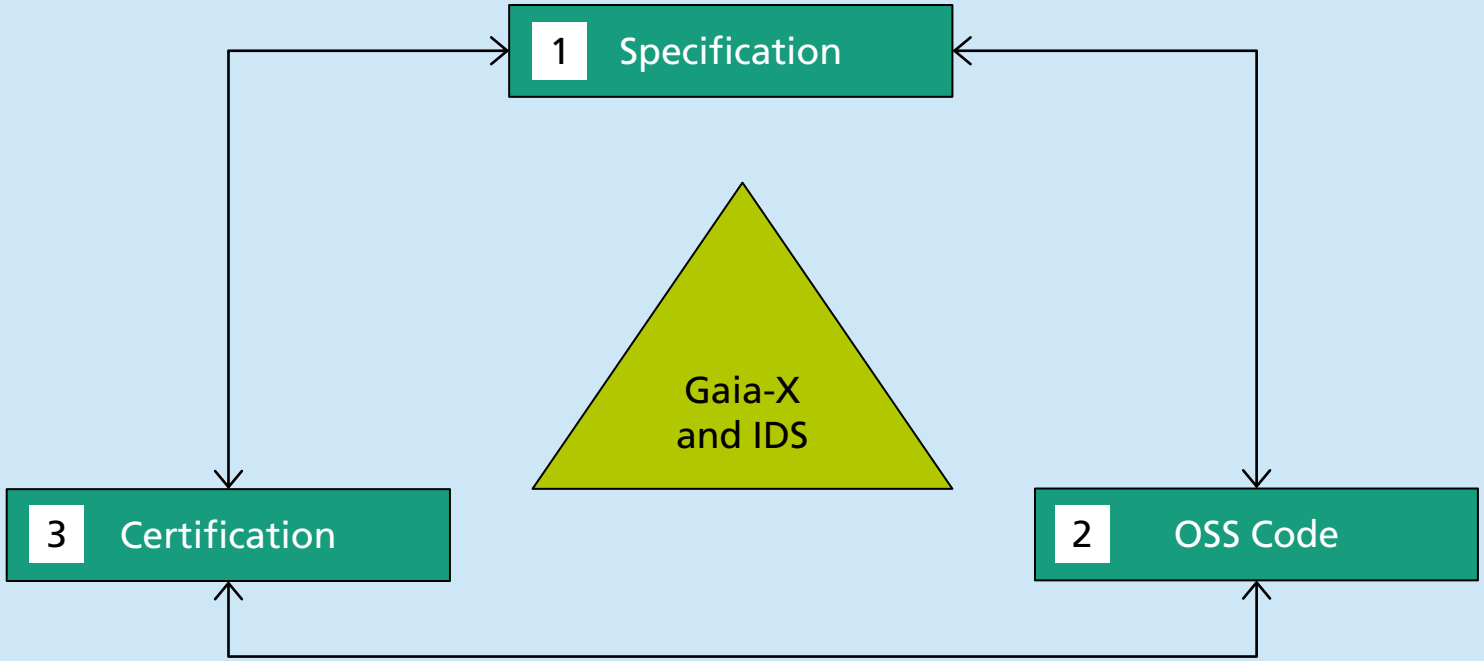
public

IDS Association and Gaia-X Association share the same goals and are closely aligned

	IDS	Gaia-X
Rationale	How to keep data sovereignty in data sharing?	How to keep sovereignty over data and services in the cloud?
Institutionalization	Non-for-profit registered association according to German law	Non-for-profit registered association according to Belgian law
Members	130+	310+
Data Space Scope	Within data spaces	Within data spaces Across data spaces
Key Deliverables	1) Specification 2) OSS code 3) Certification	1) Specification 2) OSS code 3) Labelling and Compliance
Geographical Focus	International	International w/ strong European »flavor«
Architecture Design	Federated	Federated
Key Architecture Elements	– IDS Connector – IDS Essential Services (e.g. Clearing House, Broker etc.)	– Gaia-X Federation Services
Certification Strategy	– Component Certification – Certification of Operations Environment (based on ISO 27001)	– Self-Certification based on Self-Descriptions
Identity Management Approach	– Dynamic Attribute Provisioning Service (DAPS)	– Self-Sovereign Identities

public

Both IDS and Gaia-X pursue the same »magic triangle«



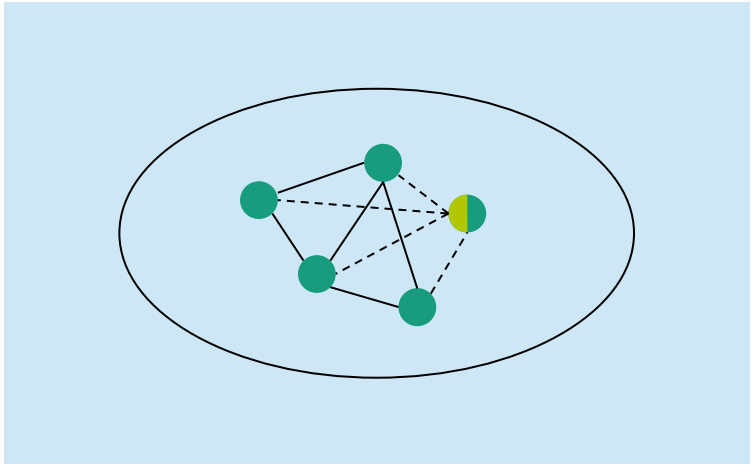
- All three angles must always be in-sync, i.e. object of a consistent release management
- DevOps principles shall be applied as soon as possible

Source: (GAIA-X AISBL, 2021).

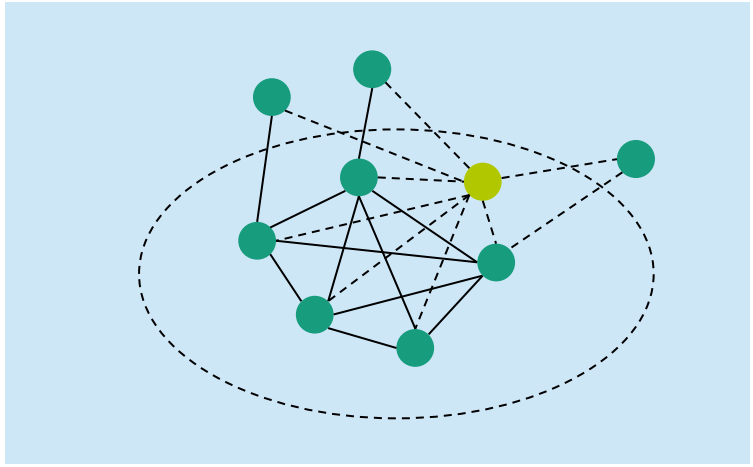
public

Data sharing ecosystems evolve over time

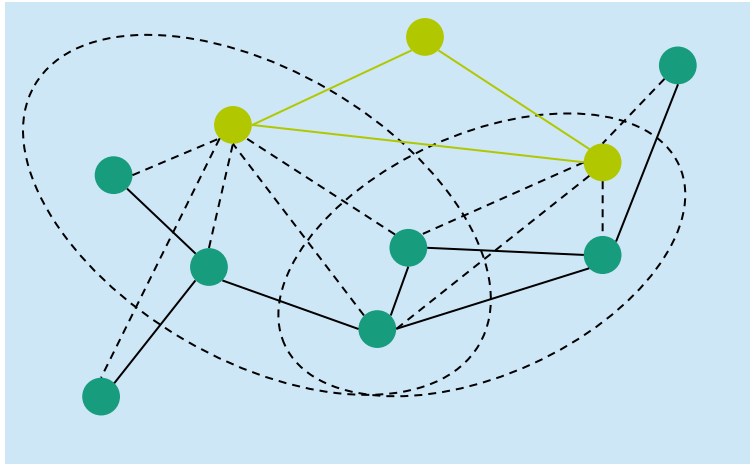
I Closed Ecosystem



II Open Ecosystem



III Federation of Ecosystems



Legend:

Roles: Participant (Data Holder | Data User); Intermediary.

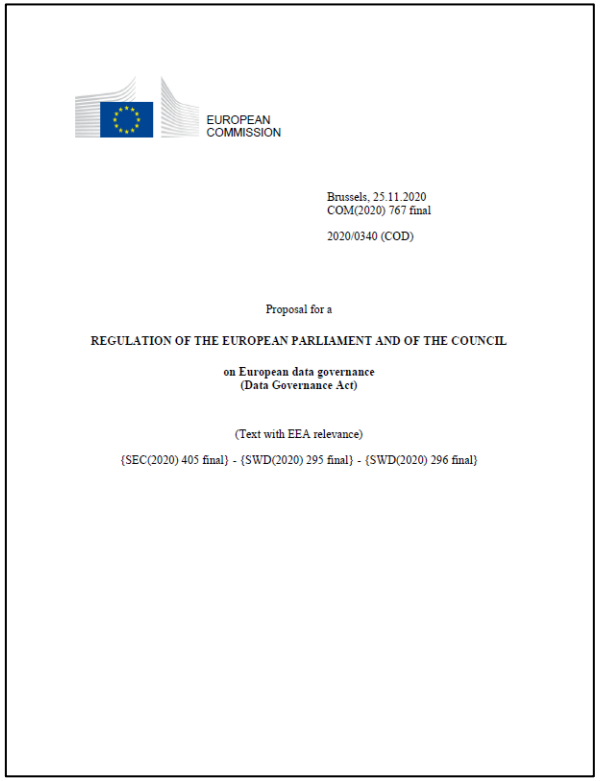
Data Exchanges: Shared data incl. metadata between participants; Metadata between participant and intermediary; Metadata between intermediaries.

Ecosystems: Closed; Open.

public

At present, there are many questions open when it comes to establishing the federator role

- Who will take over the federator role in ecosystems? The consortium of ecosystem participants? A »primus inter pares« participant? A dedicated, neutral partner?
- How does a business model for a federator look like? Not-for-profit? Low-profit? Profit-oriented?
- What implications does regulation have on the concrete instantiation of the federator role (EU Data Governance Act etc.)?



EUROPEAN
COMMISSION

Brussels, 25.11.2020
COM(2020) 767 final

2020/0340 (COD)

Proposal for a

REGULATION OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL

on European data governance
(Data Governance Act)

(Text with EEA relevance)

{SEC(2020) 405 final} - {SWD(2020) 295 final} - {SWD(2020) 296 final}

Source: EC (2020b).

public

The Data Spaces Business Alliance is good news for all data space owners

Data Spaces Business Alliance

Unleashing the Data Economy

BDV BIG DATA VALUE ASSOCIATION

FIWARE FOUNDATION

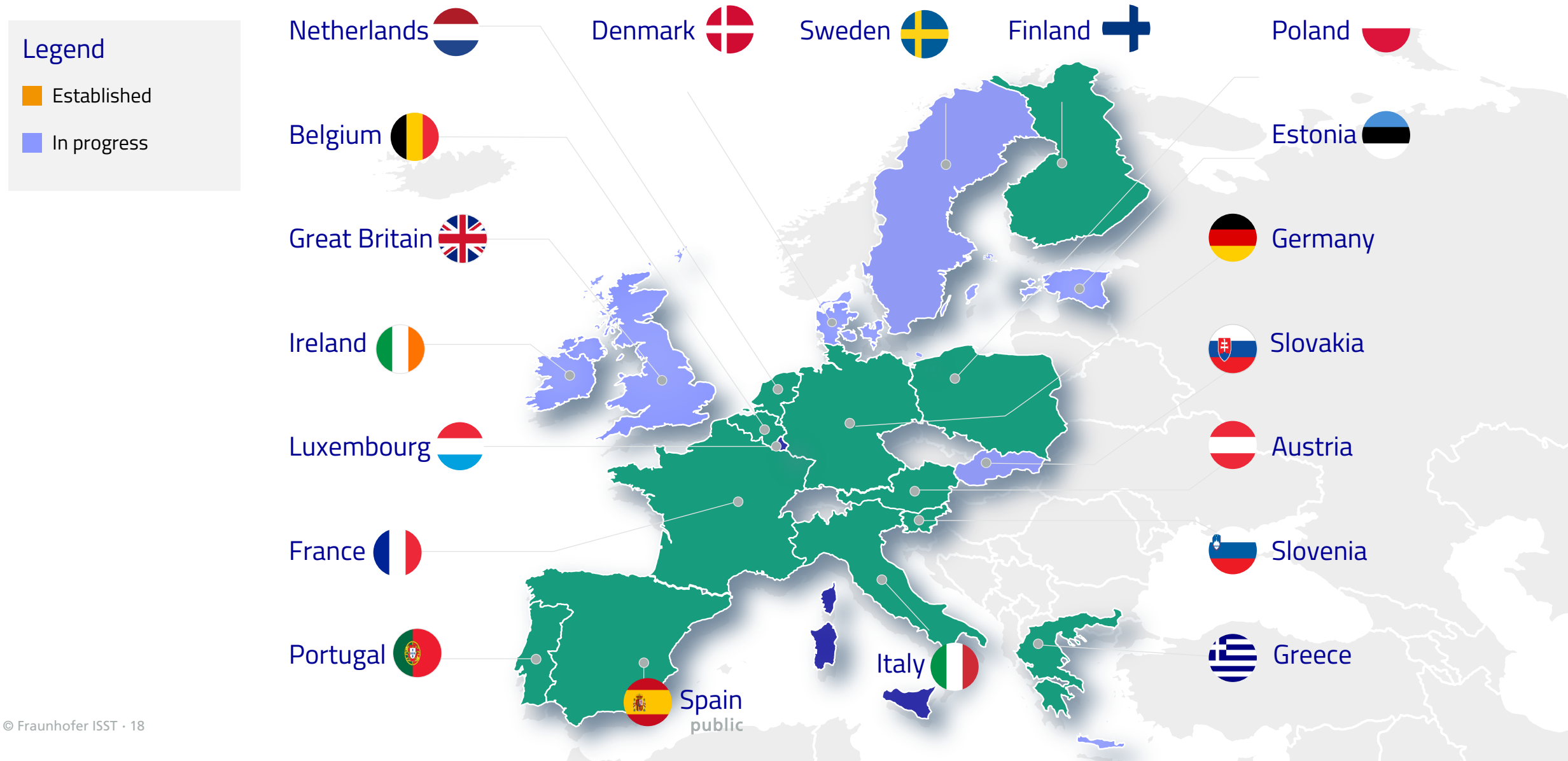
gaia-x

INTERNATIONAL DATA SPACES ASSOCIATION

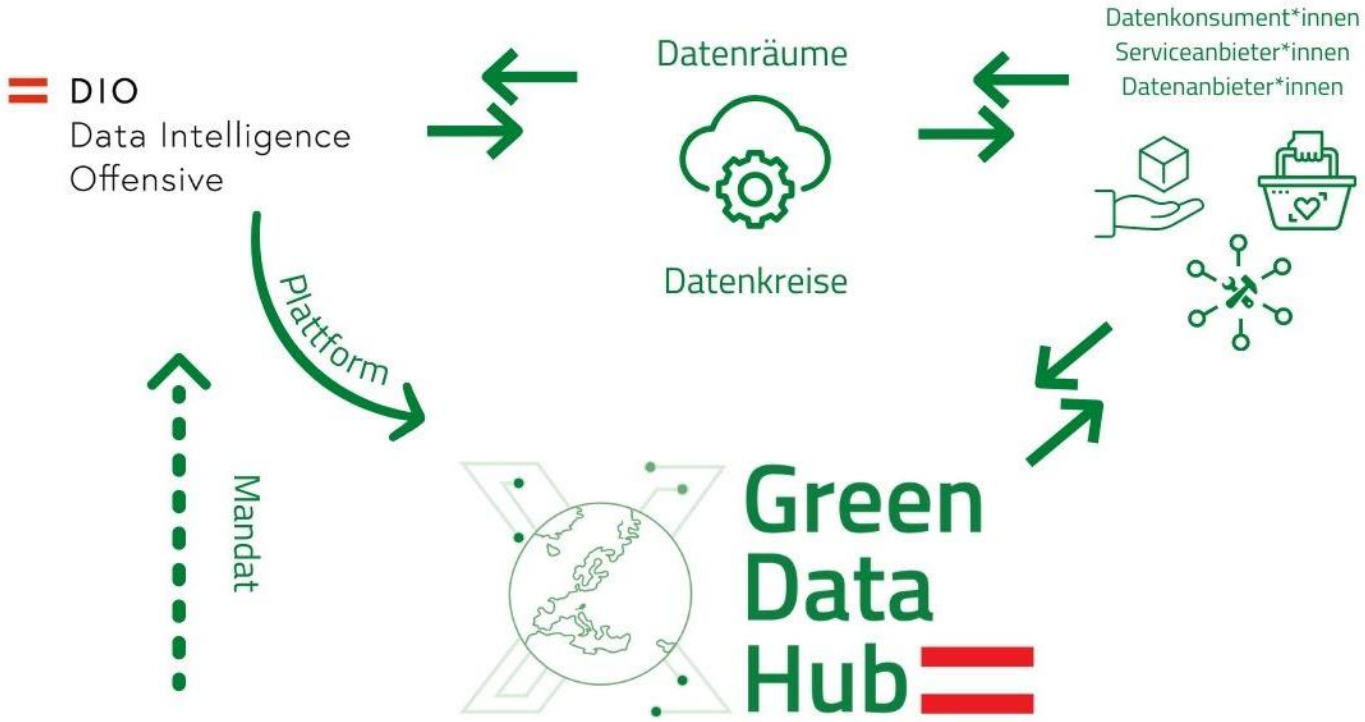
Source: EC (2020b).

public

Gaia-X is setting up a network of hubs to foster adoption of cloud and data sovereignty



Gaia-X has created a pan-European movement around data sovereignty, as the example of the Austrian Data Intelligence Offensive shows



- Gaia-X kompatible Software für
  - Data Spaces
  - Use Cases
  - Trust
  - Federation
  - Identity
  - Sovereignty
- 4 Data Spaces
  - Energiewende
  - Mobilitätswende
  - Kreislaufwirtschaft
  - Geodaten
- Kooperation
  - DE, CH
  - IDSA
  - BDVA
  - GAIA-X AISBL
  - Eclipse Foundation - EDC

Source: DIO (2022). NB: Kept in German as Original Project Language.

public

Table of Contents

public

The project aims at a Gaia-X and IDSA compliant implementation

About the Eclipse Dataspace Connector

A data-sharing system requires a protocol implementation for policy enforcement among participants. The Eclipse Dataspace Connector will implement the International Data Spaces standard (IDS) as well as relevant protocols and requirements associated with the Gaia-X and thereby provide implementation and feedback to these initatives. However, the connector will be extensible so that it can support alternative protocols.

Whatever the individual setup is — on-premises bare-metal, different cloud vendors, hybrid, even single end-user machines — the EDC can be customized to work within any environment at scale.

The connector's added value is achieved through the separation of control and data plane, which enables a modular and thereby customizable way to build dataspaces. Due to common interfaces and mapping of existing standards, the connector adds capabilities of contract negotiating and policy handling in an interoperable manner.

Open, Community-driven and extensible

As an open source project hosted by the Eclipse Foundation, the Eclipse Dataspace Connector provides a growing list of modules for many widely-deployed cloud environments (AWS, Azure, GCP, OTC etc.) "out-of-the-box" and can easily be extended for more customized environments, while avoiding any intellectual property rights (IPR) headaches.

The most important facts about the Eclipse Dataspace Connector

- The EDC is **completely FOSS** supported by various companies
- The EDC (through Eclipse Foundation) has clear and **accepted governance** structure and community processes
- The EDC is **more than connecting a database**
- The EDC manages **data transfer and flow inclusive management of contract and policy management** in cloud-native environments
- The EDC follows **a modular system** to serve as facilitator
- Running **code available** on Github (s. Developer Resources)
- **We welcome everyone** to join the community, drive the idea of dataspaces, discuss requirements, and contribute

Licenses:

Apache License, Version 2.0

Active Member Companies:

Member companies supporting this project over the last three months.

amadeus SAP Daimler TSS Microsoft Fraunhofer BMW MINI

See https://projects.eclipse.org/projects/technology.dataspaceconnector (accessed on January 31, 2022).

public

The EDC project is the strategic FOSS endeavor to achieve trusted data sharing and data sovereignty

- Associations such as Gaia-X and IDSA define the rules of the game but will not deliver ready-to-scale code – this is what the EDC project does
- The big opportunity for us as a community is to create a true FOSS de-facto standard for data sovereignty in the cloud
- It is important to have all stakeholder groups aboard, i.e. user companies, service providers, cloud platform providers, research organizations
- We need to do it now because the market for data spaces is now developing and growing fast

Let us pursue the vision that in 10 years from now nobody speaks about data sovereignty anymore.
Why? Simply because technologies such as the EDC connector are then implemented in any cloud service.

public

Thanks a lot for listing!

Univ.-Prof. Dr.-Ing. habil. Boris Otto

Fraunhofer ISST
Director
Boris.Otto@isst.fraunhofer
+49-231-97677-200

WWW.ISST.FRAUNHOFER.DE