



**ECLIPSE**  
**2021 CON**

# **Eclipse Dataspace Connector - How to build data spaces**

Markus Spiekermann  
Head of Department Data Business, Fraunhofer ISST  
Project Lead of Eclipse Dataspace Connector



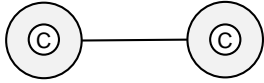
- Introduction
- Data Spaces
- Business Ecosystems and Networks
- Eclipse Dataspace Connector

# Motivation

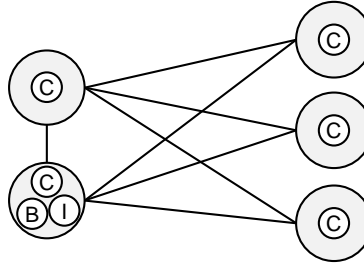


*»A real data economy, on the other hand, would be a powerful engine for innovation and new jobs. And this is why we need to secure this data for Europe and make it widely accessible. **We need common data spaces** - for example, in the energy or healthcare sectors. This will **support innovation ecosystems** in which universities, companies and researchers can access and collaborate on data. And it is why we will build a European cloud as part of NextGenerationEU - based on **GaiaX**.«*

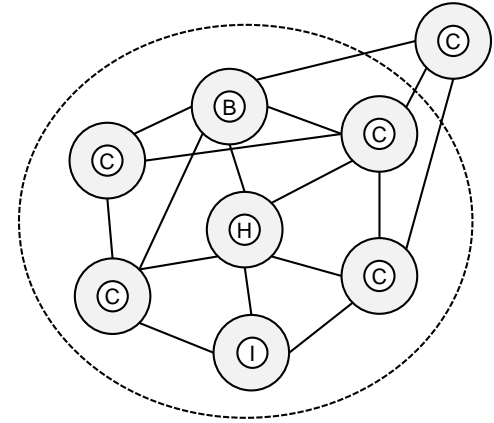
# Change of Data Sharing



Bilateral data  
exchange



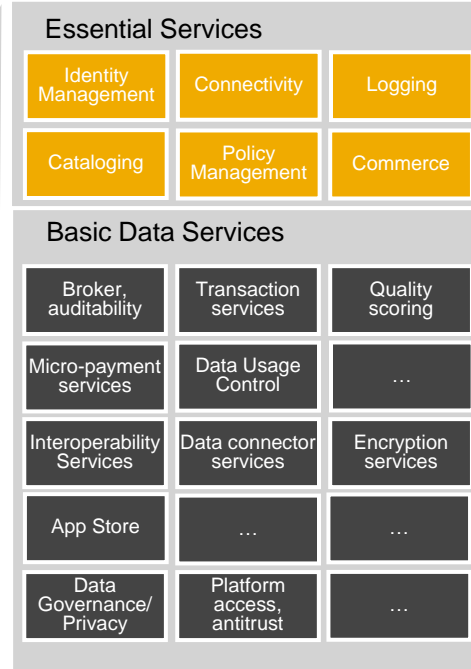
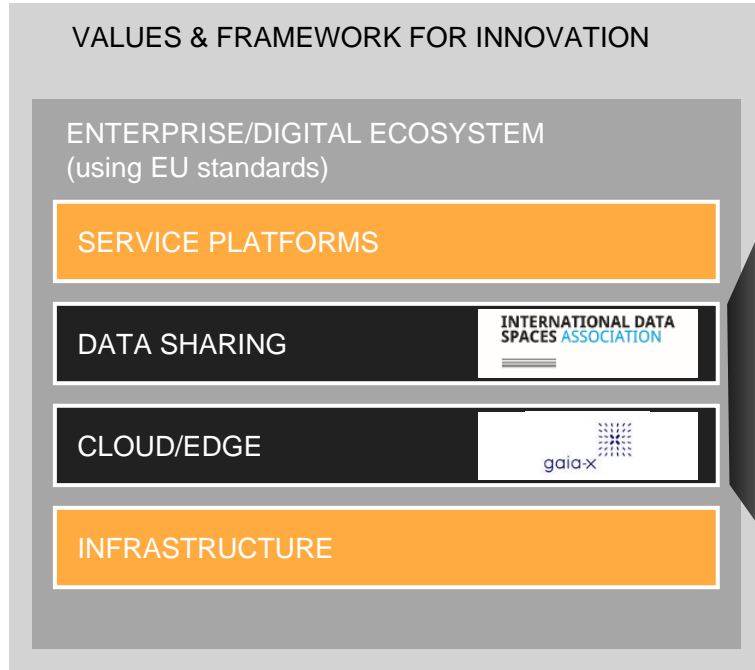
Closed group data  
exchange



Open and dynamic data  
exchange

# Data Spaces

# What does a Data Space bring to the Table?



## Design Principles

- European values
- Secure and trusted
- Easy-to-use
- Federated, neutral
- Vendor-agnostic

## Need for Action

Implement technologies and governance for data spaces that enable and ensure transparency and data sovereignty, as end-to-end control by the data provider over the use of its data across corporate boundaries.



# International Data Spaces e. V.

- Initiated: 2016
- Member: 130+ in 22 countries (09/2021)
- Architecture Document: RAM 3.0 ([link](#))
- Objectives:
  - Create a reference architecture
  - Requirement engineering in the field
  - Establish a common standard
  - Support certifiable software components

## INTERNATIONAL DATA SPACES ASSOCIATION



Identity Management	DAPS
Cataloging	IDS-Broker
Connectivity	IDS-Connector
Policy Management	IDS-Connector
Logging	IDS-Clearing House
Commerce	tbd.

# Gaia-X European Association for Data and Cloud AISBL

- Initiated: 2019
- Member: 300+ in 25 countries (09/2021)
- Architecture Document: TAD [\(link\)](#)
- Objectives:
  - Gaia-X Policies and Rules
  - Architecture of Standards
  - Federated Services



Identity Management	Federated Identity
Cataloging	Federated Catalog
Connectivity	Connector
Policy Management	Data Contact Service
Logging	Data Exchange Logging Service
Commerce	tbd.





# Business Ecosystems and Networks

# Business Ecosystems

- Cluster organizations from various interests (e.g. domain)
  - Including service provider and operating companies
- Enable collaboration for innovation and business models
- Elaborate on future requirements and challenges to be addressed
- Define common governance rules with democratic structures
- Open for new participants and technology



Images from freepik.com by tawatchai07 and aleksandarlittlewolf

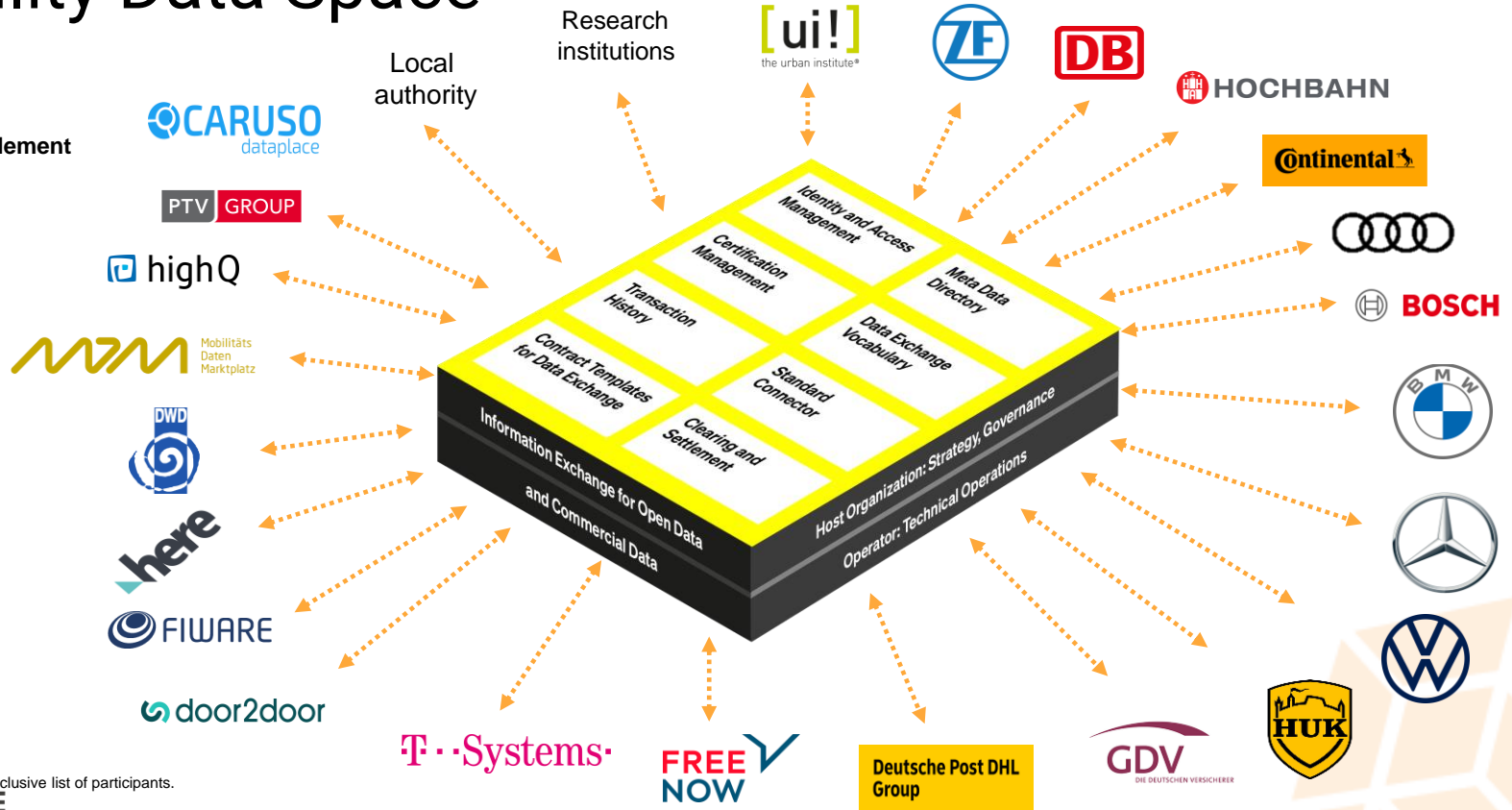
# Mobility Data Space



- Initiated by the German federal government's "Concerted Action on Mobility" committee in 2019
- Data sharing community to build the future of mobility
- Promotes forward-looking mobility services
- Adapt the IDS RAM and available open source implementations of components
- More than 200 stakeholders of German mobility landscape, science, business and government
- 20+ use cases that are presented this week on the ITS Worldcongress
- Productive operation planned from early 2022

# Mobility Data Space

Data sharing enablement  
(bi-directional)



NB: Viewgraph shows non-exclusive list of participants.

# Catena-X Automotive Network



- Founding of Catena-X Automotive Network e.V. took place on 07.05.2021.
- Alliance for secure and standardized data exchange along the automotive value chain
- Offer network and technologies for collaboration and innovation
- Ensure the economic viability of all network partners
- Connect to cross-industry networks “built a Gaia-X compliant IDS-system”
- Technical components and services incl. transfer and scale out
- Initial use cases, e.g. Traceability, CO<sub>2</sub> Footprint, Circular Economy, Demand and Capacity Mgt.
- Consortium of Industry, technology and platform experts

# Catena-X Automotive Network

## Data Sovereignty & Interoperability (European Architecture)



Decentralized Data Spaces



Competitive at Application Level

## One Operating System (decentralized, federated)



Operating System on GitHub



Plug and Play - Standardized APIs

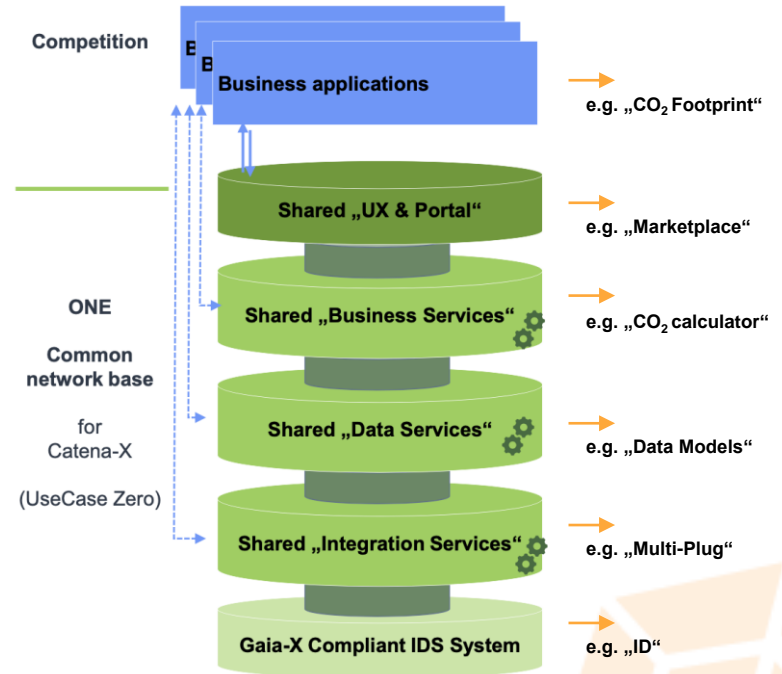
## Collaborative und Agile Product Development



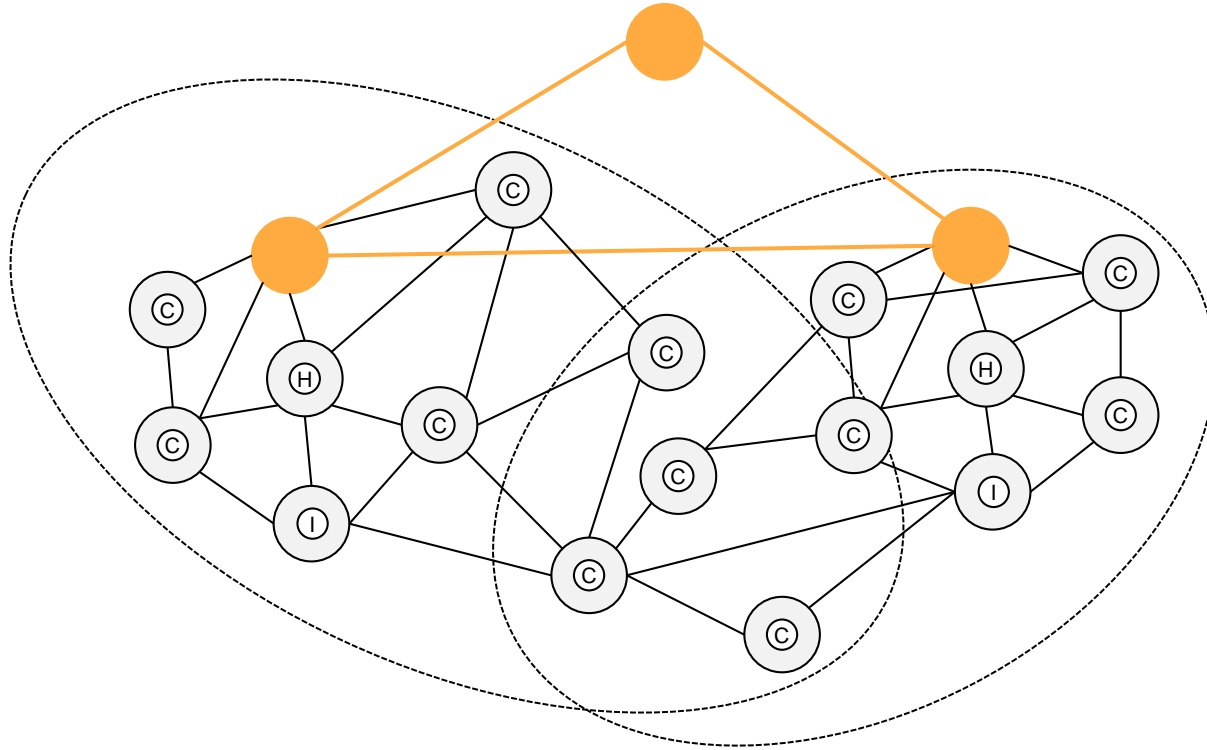
Eclipse Open Source Community



100% Agile Working Model



# Implications of Ecosystems and Data Spaces



# Challenge accepted

- Catena-X, Mobility Data Space, and other project use cases present new data sharing challenges
  - Connection with multiple data spaces
  - Support for data flow and transfer protocols to handle diverse data types
    - Streaming, Big Data
  - Identity across multiple jurisdictions
  - Cataloging across many providers
  - Policy management that traverses multiple infrastructure layers
  - Need for extensibility and modularity to accommodate diverse needs and use cases



# Eclipse Dataspace Connector

# Eclipse Dataspace Connector

[Overview](#)[Downloads](#)[Who's Involved](#)[Developer Resources](#)[Governance](#)[Contact Us](#)

Existing open-source projects address the technical challenges of cataloguing and transferring data for a wide range of use cases. However, there is no open-source effort aimed at providing an interoperable, cross-organization framework for data sharing that is built on a common identity model and uniform policy enforcement. This project will integrate with existing data exchange technologies and provide these missing pieces to create a system for data sharing where each organization is able to exert control over how its shared data is used.

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 associated with the GAIA-X project. However, the connector will be extensible so that it can support alternative protocols.

This project will provide implementation and use case feedback to IDS and GAIA-X.



Markus Spiekermann  
Fraunhofer ISST



Alexandru Danciu  
SAP



Stefan Ettl  
BMW Group



Moritz Keppler  
Daimler TSS



Florian Seidel  
Amazon AWS



Julia Pampus  
Fraunhofer ISST



Werner Jost  
Deutsche Telekom



Franco Wolf  
ZF Friedrichshafen



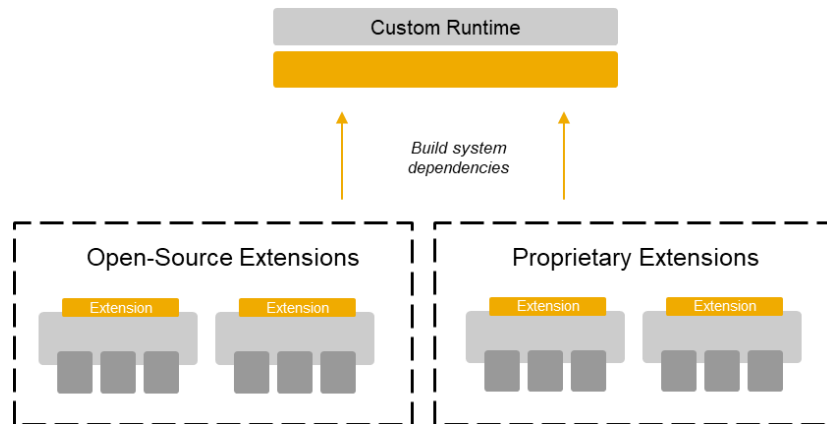
Jim Marino  
Microsoft

# Features

- Based on a simple modularity system
- Separate control and data planes
- System is asynchronous and highly available
- Transfer processes are fully auditable
- Eliminate single points of failure
- Cloud aware policy enforcement and projection
- System security

# The Connector is a Coordinator

- Modules are assembled into a runtime
- Data cataloging handled by external system
  - e.g., Apache Atlas
  - Modularity allows this to be substituted
- Data transfer and storage performed out of band
  - Modularity allows this to be substituted or augmented
    - Ability to add multiple transfer mechanisms to support diverse data types
  - Solves key aspects of how to handle streaming, large data transfer, and hyper scaling

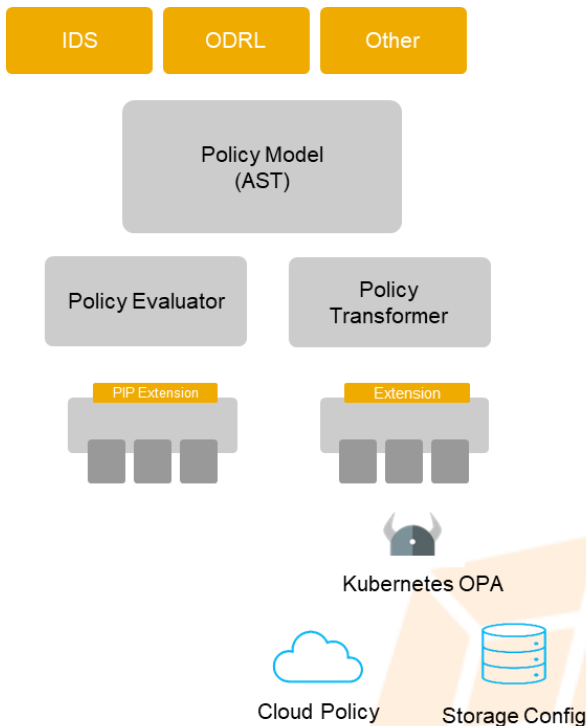


# The Connector is a Coordinator

- Processes are modeled as persistent state machines
  - Nearly symmetric on the client connector and provider connector
  - Includes states for initiation, receipt, resource provisioning, transfer, and termination
    - State transitions happen asynchronously after a defined task is completed
      - For example, provisioning cloud storage for data
    - Tasks must be idempotent for reliability
  - Processes are persisted to a store
    - Can be a highly available database with geographic fail over
    - Or, as simple as an in memory Map
- Provides a foundation for full observability and data audit
  - Metadata about each state is recorded
  - Observability can be achieved by correlating the process id with separate telemetry data

# Cloud-Aware Policy Enforcement

- Policy engine that is cloud aware
  - Not limited to the connector
  - Can be embedded in other processes
- Horizontal and vertical enforcement
- Parses policy syntax into an internal Policy Model AST
- Evaluators and transformers to enforce policy
  - Evaluators can make policy decisions, e.g. is a connector authorized
  - Transformers can create and deploy policy to different levels
    - OPA, storage, etc.



# Status Quo

- GitHub: <https://projects.eclipse.org/proposals/eclipse-dataspace-connector>
- Official Eclipse project since June 2021
- Initial contribution has been made
  - Basis for data exchange and connectivity available
  - IDS protocols are currently merged
- Expand the community
  - Onboarding „Experience“ on YT:  
<https://www.youtube.com/channel/UCYmjEHtMSzycheBB4AelTHg>

# Thank you!

Join the conversation:



[@EclipseCon](https://twitter.com/EclipseCon) | [#EclipseCon](https://twitter.com/EclipseCon)



**ECLIPSE**  
**2021 CON**





**ECLIPSE**  
**2021 CON**

# **Eclipse Dataspace Connector - How to build data spaces**

Markus Spiekermann  
Head of Department Data Business, Fraunhofer ISST  
Project Lead of Eclipse Dataspace Connector

