



Eclipse Dataspace Connector – How to build Data Spaces

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

- *Introduction*
- *Business Ecosystems*
- *Data Spaces*
- *Eclipse Dataspace Connector*
- *Relation to IDS(A)*
- *Relation to Gaia-X*
- *Key Takeaways & FAQ*

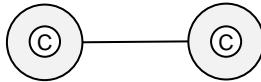


Motivation

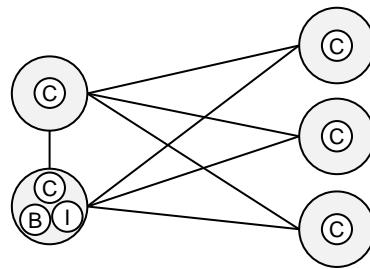


»A real data economy, on the other hand, would be a powerful engine for innovation and new jobs. And this is why we need to secure this data for Europe and make it widely accessible. **We need common data spaces** - for example, in the energy or healthcare sectors. This will **support innovation ecosystems** 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**.«

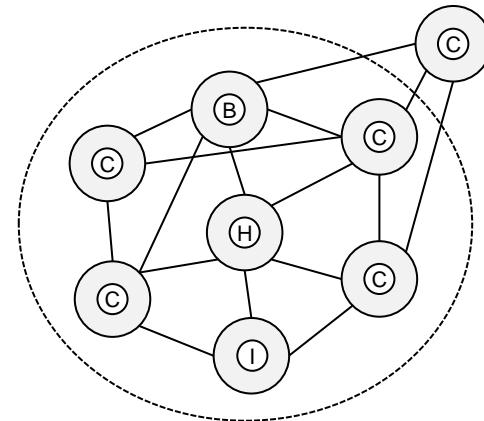
Change of Data Exchange and Sharing



Bilateral data exchange



Closed group data exchange



Open and dynamic data exchange

Business Ecosystems

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*
- *Require openness for new participants and technology*
 - *Adapts the concepts of data spaces and their reference architectures*



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, CO₂ Footprint, Circular Economy, Demand and Capacity Mgt.*
- *Consortium of Industry, technology and platform experts*

Mobility Data Space

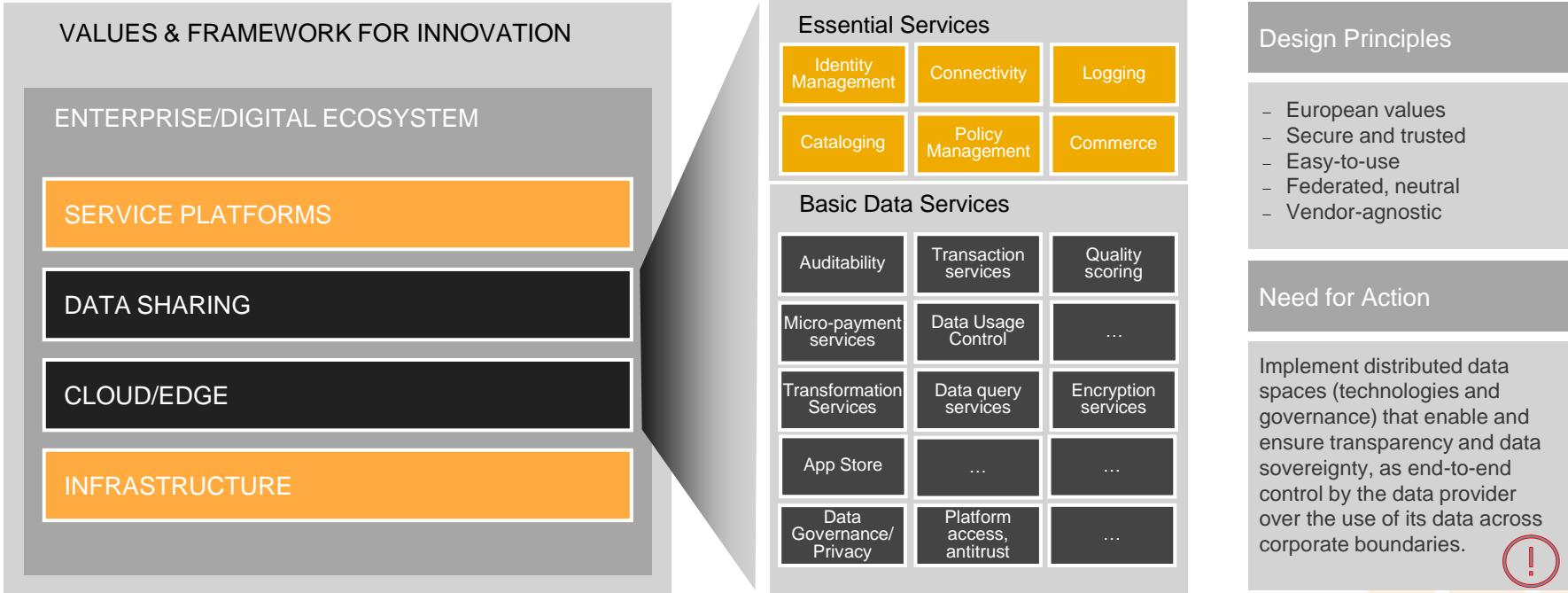


Mobility
Data Space

- *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 were presented on the ITS Worldcongress 2021*
- *Productive operation planed from early 2022*

Data Spaces

What does a Data Space bring to the Table?

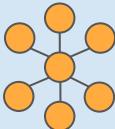
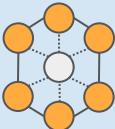
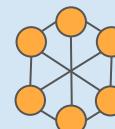


Dataspace Characteristics

- A *dataspace is a way for organizations to securely share data with other participants.*
- *Dataspaces are built on identity, trust, policy, and interoperability*
- *Dataspaces enable data cooperation in a multi-cloud federation*
- *Connections in a dataspace are always peer-to-peer*
- *Multiple participants can cooperate, but data is always exchanged 1:1*
- *Dataspaces enable a multi-cloud federated environment*
 - *On Premises, Edge, Hyperscale Cloud, Multi-Cloud*
- *Participants can have multiple roles*
 - *Data Owner, Data Holder, Data Processor, Data Recipient, Algorithm Provider,...*



Architectual Pattern

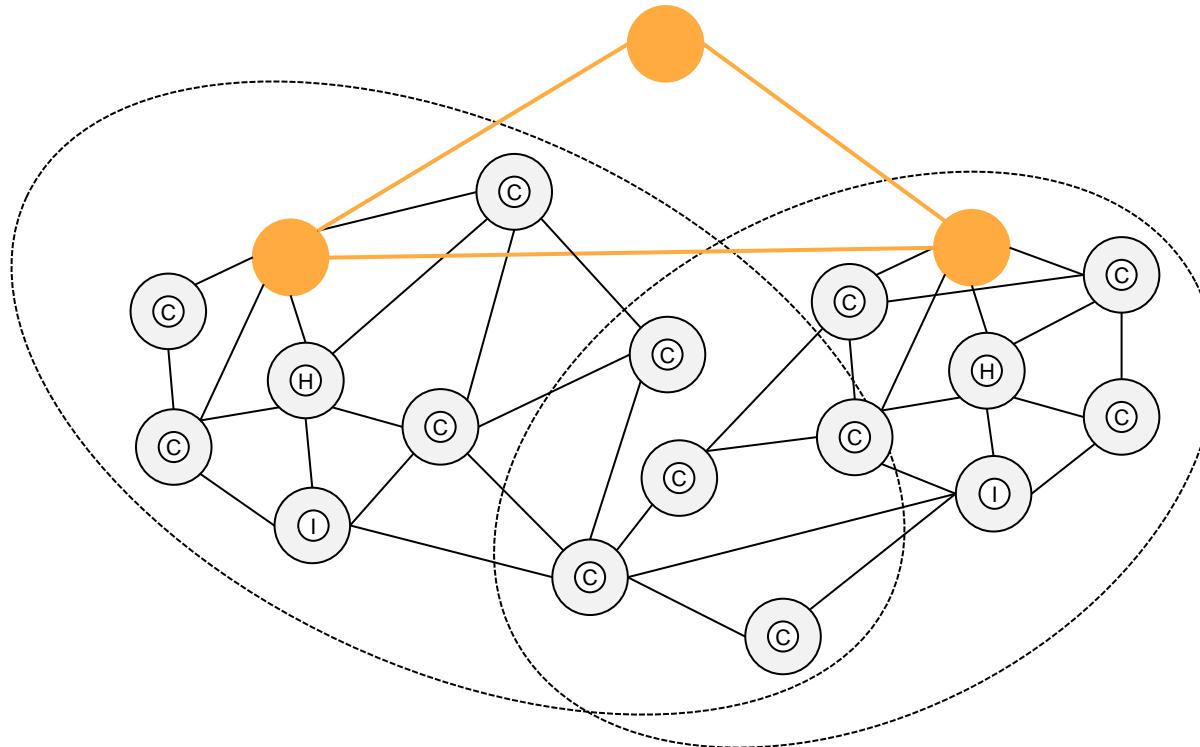
	Central Architectures	Federated Architectures	Distributed Architectures
			
Data Ownership	Central or distributed	Distributed	Distributed
Data Stewardship	Central or distributed	Distributed	Distributed
Data Capture and Creation	Distributed	Distributed	Distributed
Data Storage	Central	Distributed	Distributed
Data Enrichment and Data Preprocessing	Central	Distributed	Distributed
Data Sovereignty	Central (if any)	Distributed	Distributed
Data Provenance	Central (if any)	Central	Distributed
Data Brokering, Clearing, Billing	Central	Central	Distributed

New challenges arise!

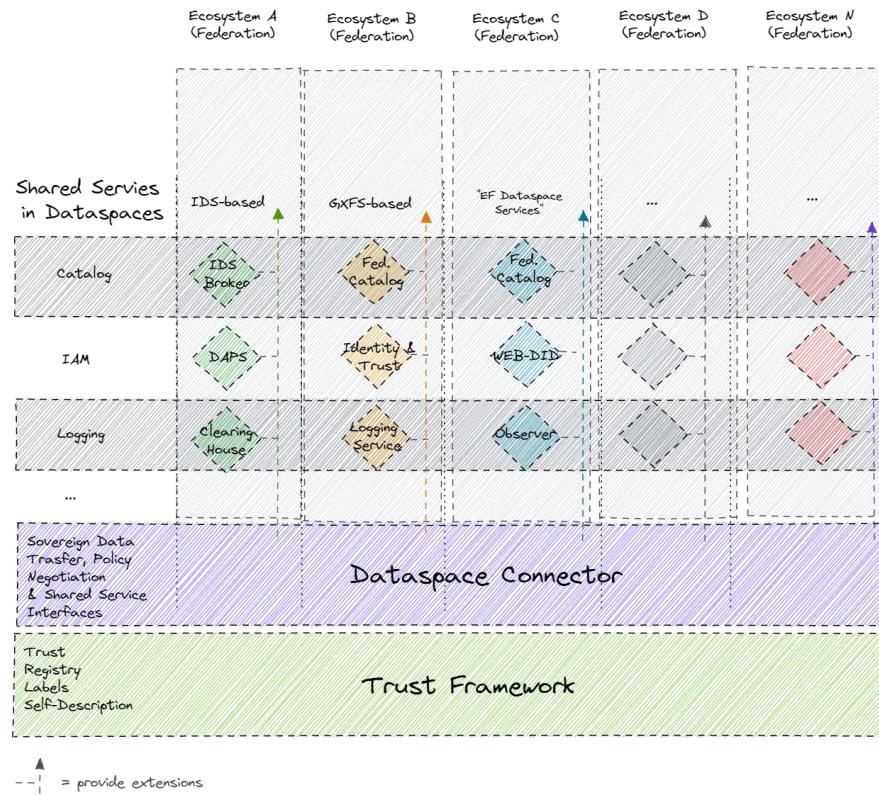
- *Different architectures and implementations for data space services*
 - *Organizations participate in various data spaces*
 - *Connection and interoperability with multiple data spaces*
 - *Identity across multiple jurisdictions*
- *Catena-X, Eona-X, Mobility Data Space, and other initiatives present new data sharing challenges*
 - *Support for data flow and transfer protocols to handle diverse data types*
 - *Push, Streaming, Large Volumes, Realtime*
 - *Cataloging across many providers*
 - *Policy management that traverses multiple infrastructure layers*
 - *Need for extensibility and modularity to accommodate diverse needs and use cases*



Implications of Ecosystems and Data Spaces



Big Picture



Eclipse Dataspace Connector

Eclipse Dataspace Connector

- *In order to build up and participate in a dataspace it's not enough to consider existing data transfer protocols and a corresponding '**data plane**'. A common standard is needed for the '**control plane**': i.e., for discovering, connecting, (automated) contract negotiation, policy management, or auditing. Dataspace connectors act as logical gatekeepers that sit within each participant's infrastructure and communicate with each other.*
- *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.*



Eclipse Dataspace Connector

- *Origin in the development of the Dataspace Connector*
 - *Limited to IDS-based data protocol*
 - *Undefined contribution guidelines and IP rights*
- *Founded in June 2021 as Eclipse Foundation project*
 - *Clear governance structures and rules for contributions based on meritocratic principles*
 - *Well known and whitelisted OSS-license with Apache 2.0*
 - *Open for everyone and all kind of contribution*
- *Growing community of partners*
 - *Amadeus, AWS, BMW, DIO, Fraunhofer, Mercedes-Benz, Microsoft, SAP, T-Systems, ZF*
 - *34 Contributors to the code-repository*
 - *Work (discussion & implementation) takes place transparently on Github*



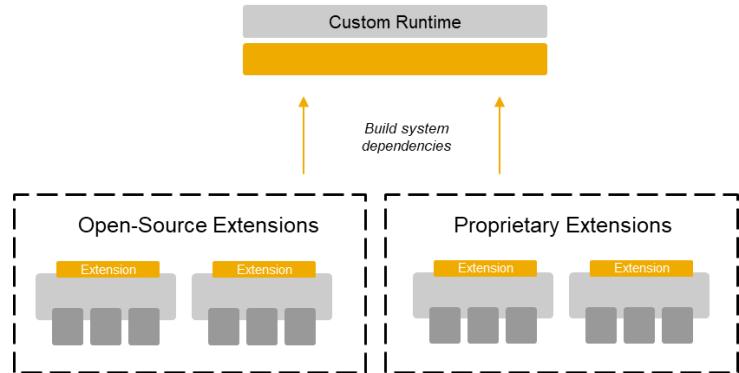
Status-Quo

- *Comprehensive code-base available on Github*
 - <https://github.com/eclipse-dataspaceconnector/DataSpaceConnector>
 - *Core-features and extension*
 - *Insights and onboarding through step-by-step samples*
- *Excerpt of available features*
 - *Identity Management via DAPS or DID*
 - *DataAsset management*
 - *Data transfer (i.a. IDS-Multipart)*
 - *Policy Management (ODRL)*
 - *Contract negotiation*



Architectural Principles

- *Modular*
 - *All functionality is contributed as a module*
 - *Lightweight, composable runtime*
 - *Minimal dependencies*
- *Extensible*
 - *Defines extension points for all features*
 - *Swap implementations, e.g., database, security*
 - *Create your own features and capabilities*
- *Adaptable*
 - *Deploy to diverse environments*
 - *Scale up and down*
- *Resilient*
 - *Leverages existing infrastructure*



Milestones and Cadence

- *Milestones planned for ~6 to 8 weeks*
 - *Future milestones describe an outlook on features to be considered*
 - *Milestones are subject to change due to the nature of OSS and agile development*
- *Triage of Features and Issues based on user requirements and sponsors*
 - *Happens close to start of next milestone*
 - *Everyone is able to influence priorities and address features*
- *Transparent process on Github which enables participation*
- [*https://github.com/eclipse-dataspaceconnector/DataSpaceConnector/milestones*](https://github.com/eclipse-dataspaceconnector/DataSpaceConnector/milestones)

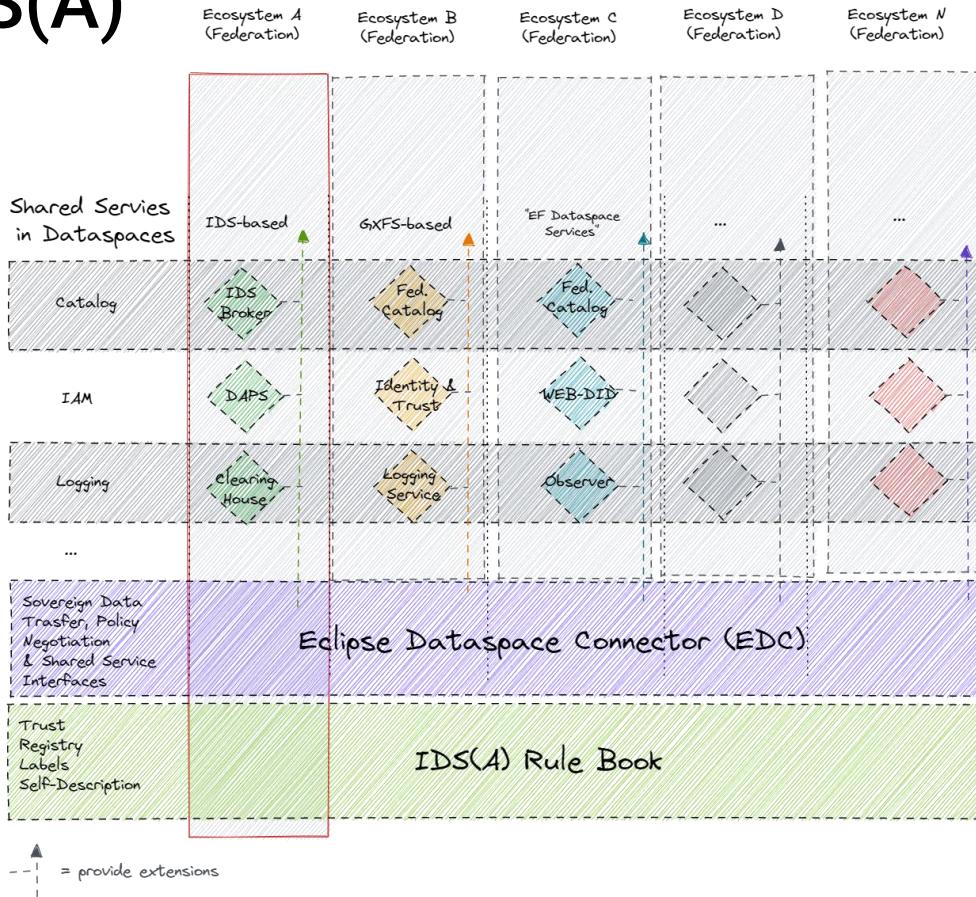
Relation to $\text{IDS}(A)$

Relation to IDSA

- *Support development of IDSA Reference Architecture Model 4.0*
- *Part of the IDSA Open Source Landscape*
- *Participation in IDSA WG Architecture and Rule Book*
- *Already supporting IDS-based messages and policy definitions*
- *Support of IDS-Self Description*
- *Support of essential services (DAPS)*
- *Maintaining IDS-Serializer and CodeGen tools*



Relation to IDS(A)



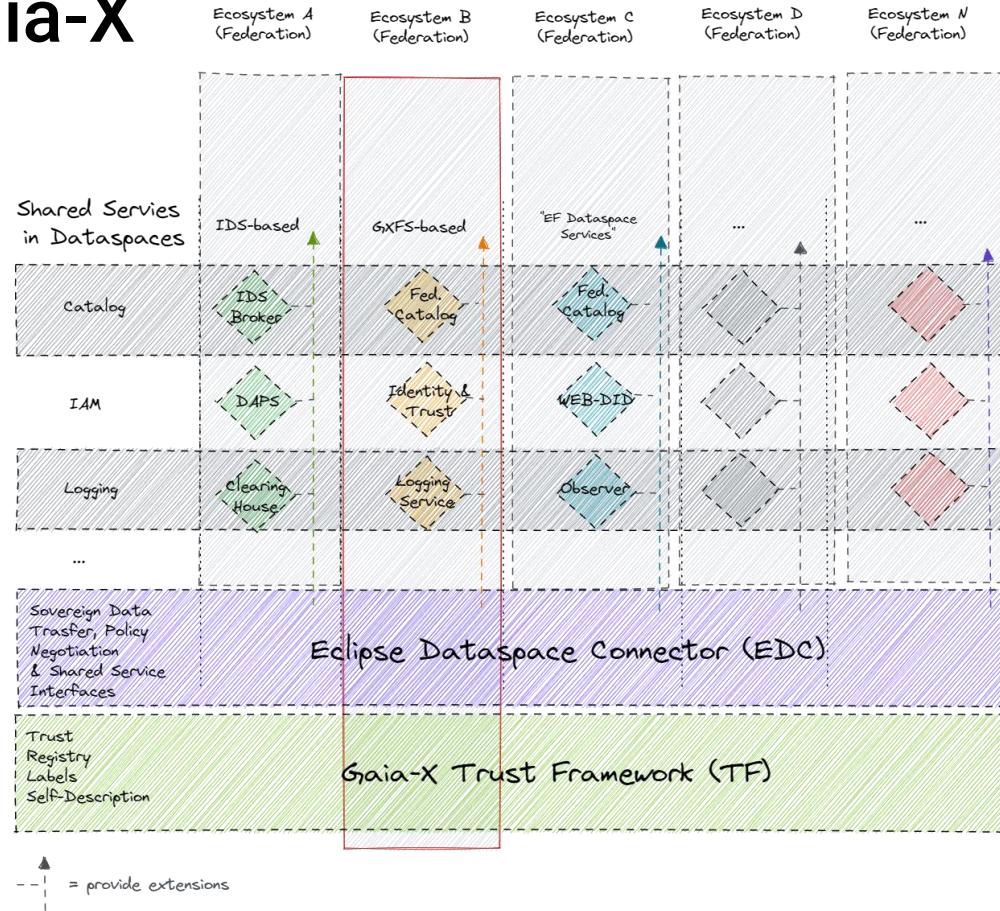
Relation to Gaia-X

Relation to Gaia-X

- *Actively contribute & support in Gaia-X WG*
- *Fulfillment of the mandatory and further criteria*
 - *Support for Gaia-X Trust Framework*
 - *Support Gaia-X compliant Self-Description*
 - *Gaia-X Registry Extension and support for VC*
- *Possible Integration with GXFS-DE implementation project*
 - *Evaluated existing specifications*
 - *Waiting for tangible code to evaluate integration*
- *Alignment with lighthouse projects to meet their requirements for building economically viable dataspaces and using Gaia-X*



Relation to Gaia-X



Key takeaways & FAQ

Key takeaways

- *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 for shared services*
- *Code and discussions are completely open available on Github*
- *We welcome everyone to drive the idea, contribute, and grow the community*

Further information and contact

General project information:

- <https://projects.eclipse.org/projects/technology.dataspaceconnector>

Github-Repository:

- <https://github.com/eclipse-dataspaceconnector/DataSpaceConnector>

Mailing list:

- dsconnector-dev@eclipse.org

YT-Channel

- <https://www.youtube.com/channel/UCYmjEHtMSzycheBB4AeITHg>



FAQ

- *Who owns the EDC project?*
- *Why is there so little documentation?*
- *Is EDC already a product?*
- *Is EDC providing Identities and Catalog capabilities?*
- *How can I contribute?*

