



TEADAL



ENABLING DATA SHARING BY EXTENDING THE DATA MESH IN A FEDERATED ENVIRONMENT

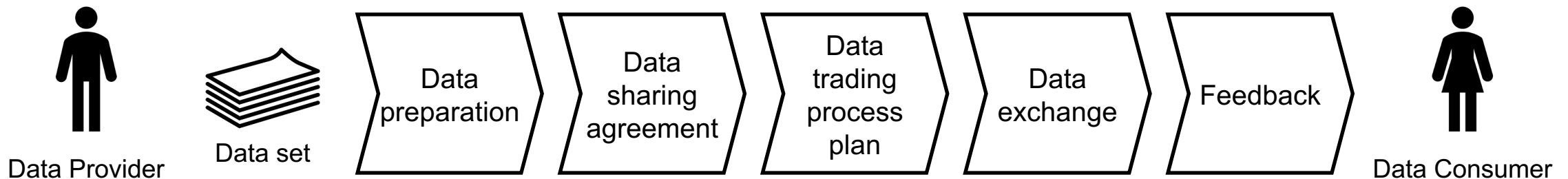
Pierluigi PLEBANI
Politecnico di Milano

TEADAL 1st Workshop, Milan, Italy
Milan, 14 March 2024

WWW.TEADAL.EU

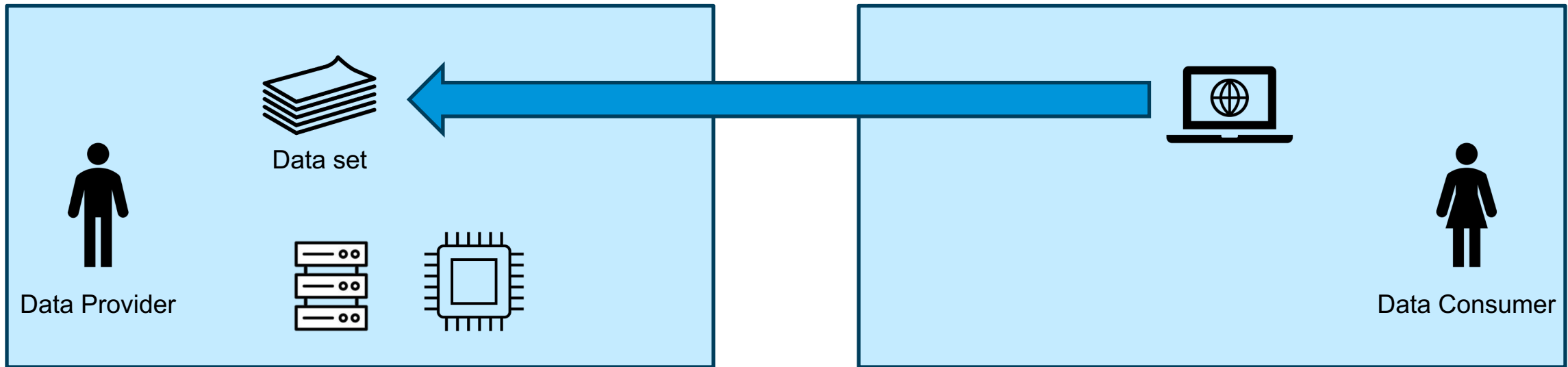
Data sharing

- Data sharing is the domain-independent process of giving third parties **access** to the data sets of others. [...]
- What the data may be used for and how it is made available is determined within the framework of the (legal) **agreements** [...]



Jussen, I., Schweihoff, J., Dahms, V., Möller, F., and Otto, B. (2023): Data Sharing Fundamentals: Definition and Characteristics. In Proceedings of the 56th Hawaii International Conference on System Sciences (HICSS), Maui, Hawaii, USA.

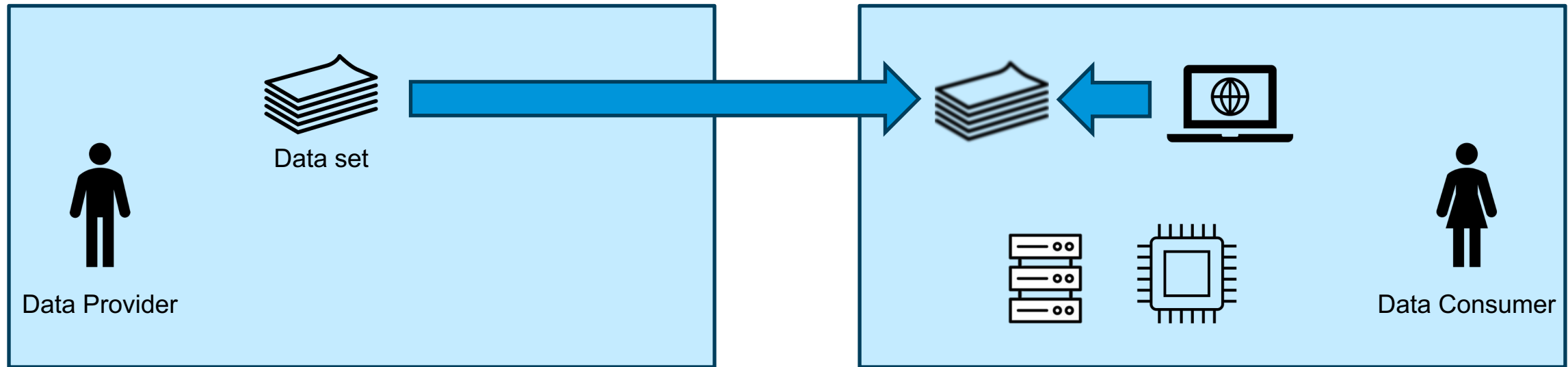
Data access: continuous access / move application



Data is fully controlled by the data provider
X Scalability required

X Data are far-away from the application
Limited resources are required

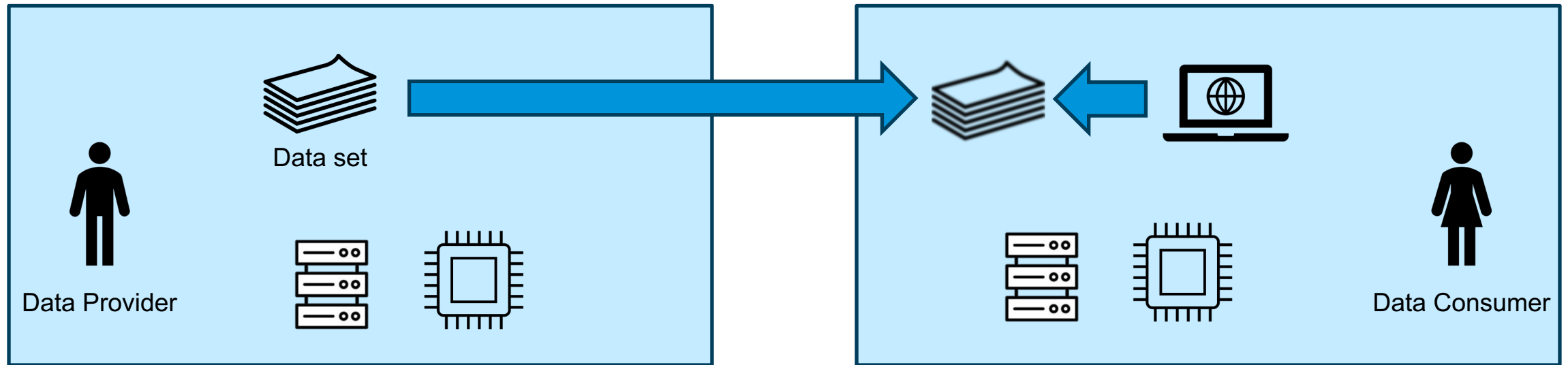
Data access: move data



- X Tailored copy of the data is needed for each consumer
- X Full control is not ensured
- Limited resources are required

- Latency is reduced
- X Computational resources required

Data access: mixed environment



Critical computation at provider side
Limited scalability
(full) control on the data

Latency is reduced
Limited resources are required

Data sharing agreement

- Data sovereignty (data provider):
 - Be sure that business critical data are not disclosed
 - Be sure that shared data are properly used
 - Being in the (full) control of my data
- Data trustworthiness (data consumer):
 - Data have quality
 - Data are reliable

Data mesh in a nutshell



Socio-technical system that can support organizations in better managing data for analytical purposes

Domain ownership

- Responsibilities of the data are given to the people that are closer to them
- Talking about people and not technology

Data as a product

- having in mind the final consumer
- data need to be curated, properly described, made visible, and easily and efficiently accessible

Self-service data platform

- a common platform offering a set of capabilities to support the data life-cycle managed is offered

Federated computational governance

- common data governance based on policies that enactment needs to be automated as much as possible

Data mesh and trustworthiness

Socio-technical system that can support organizations in better managing data for analytical purposes

Domain ownership

- Responsibilities of the data are given to the people that are closer to them
- Talking about people and not technology

Data as a product

- having in mind the final consumer
- data need to be curated, properly described, made visible, and easily and efficiently accessible

Self-service data platform

- a common platform offering a set of capabilities to support the data life-cycle managed is offered

Federated computational governance

- common data governance based on policies that enactment needs to be automated as much as possible

Data mesh and sovereignty



Data product owner is not the data owner

Sovereignty is out of the scope of data mesh

Data mesh is about a **single** organization where internal teams are federated

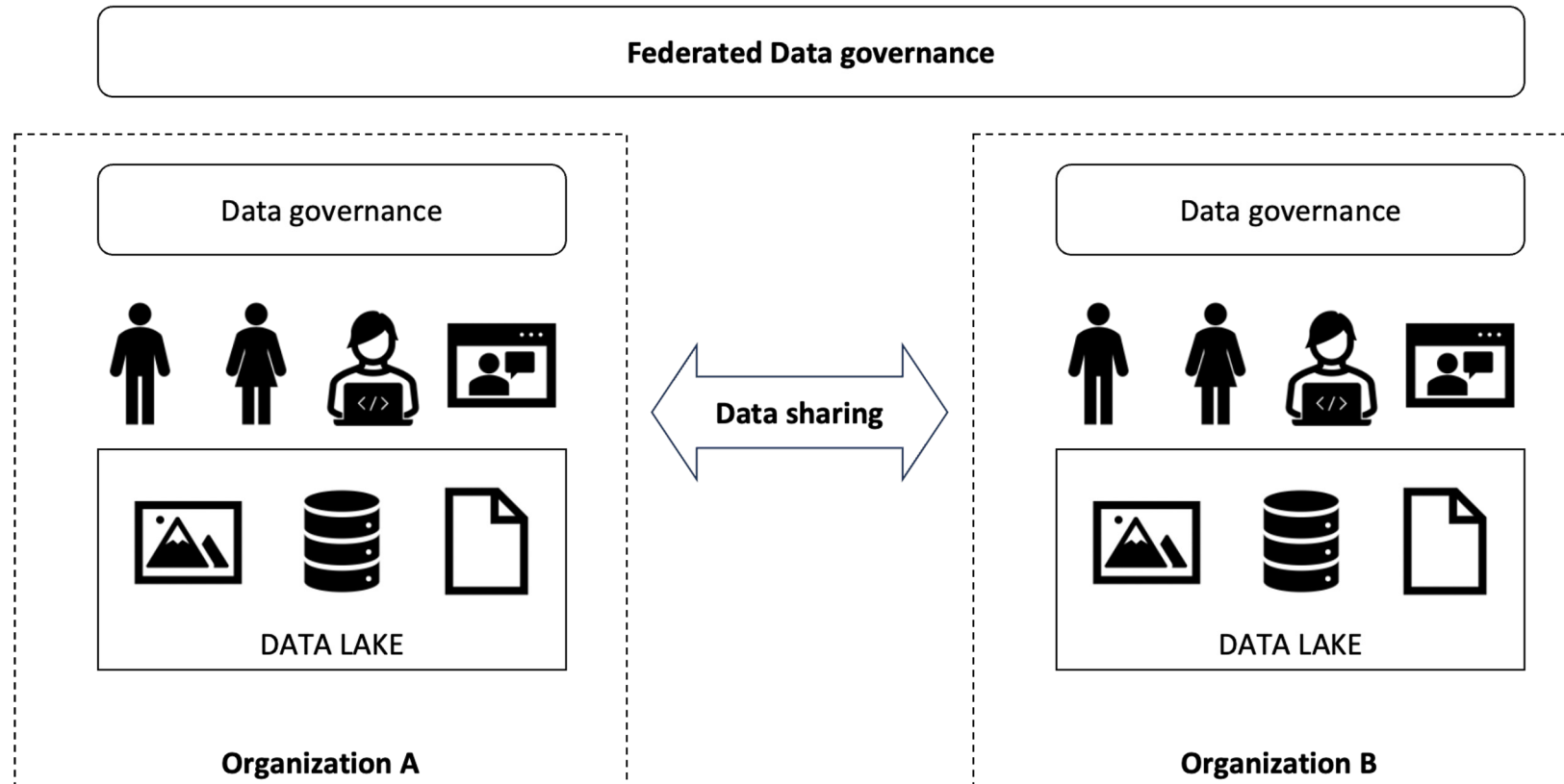
TEADAL is about a **federation of organizations**

- We want to improve the single node (via stretched data lakes)
- We want to improve the trustworthiness (via a trusted federation)

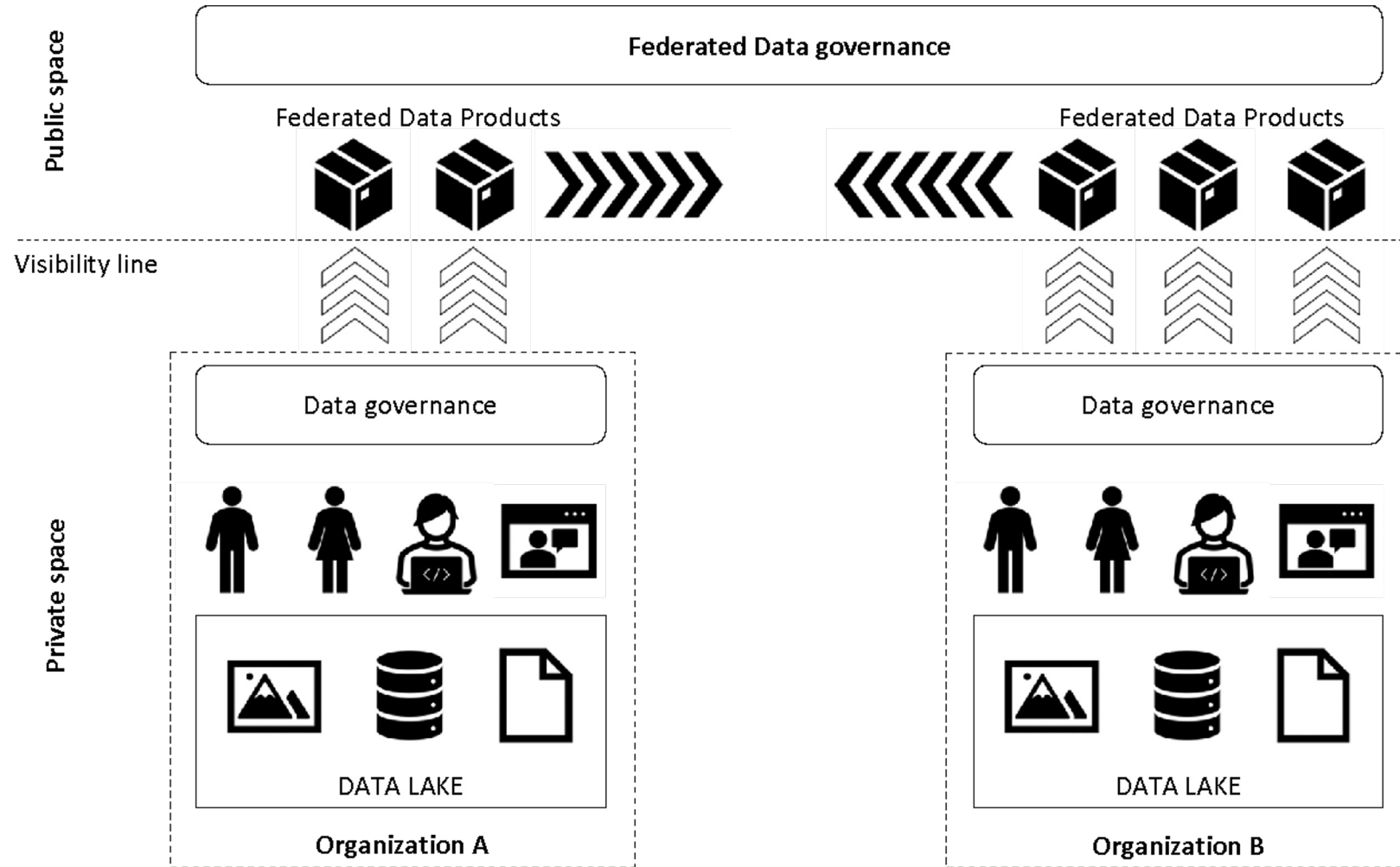
Main questions:

- Are the data mesh principles applicable to **federation of organizations** that want to share data?
- Which are the implications in this choice?
- Which are the tools that could be offered to support this vision?

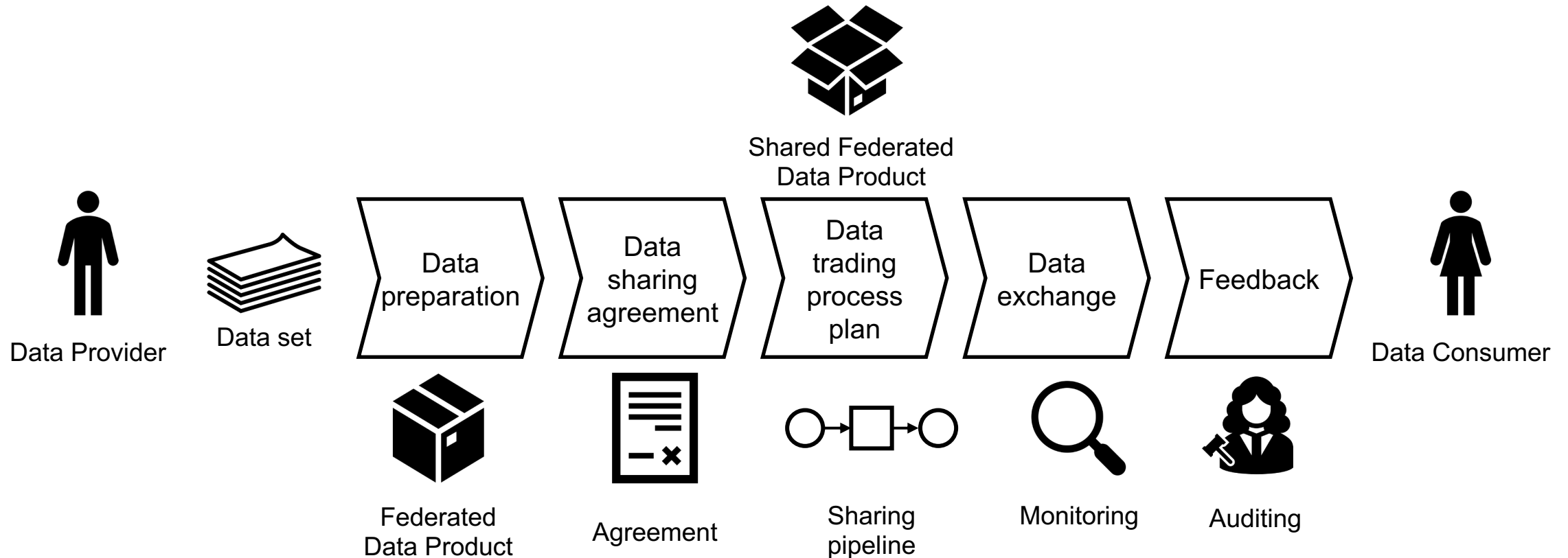
TEADAL approach



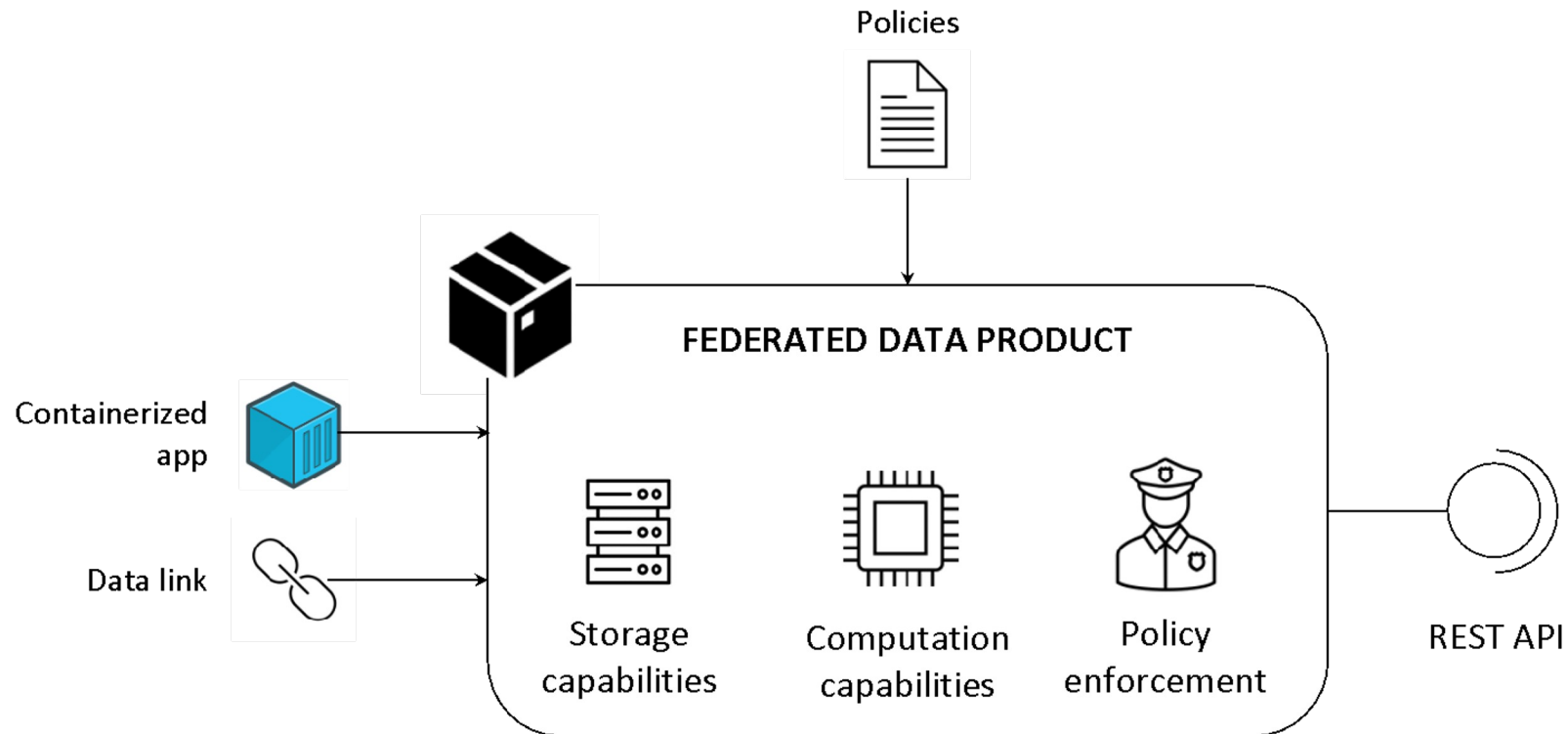
Applying Data mesh principles to TEADAL



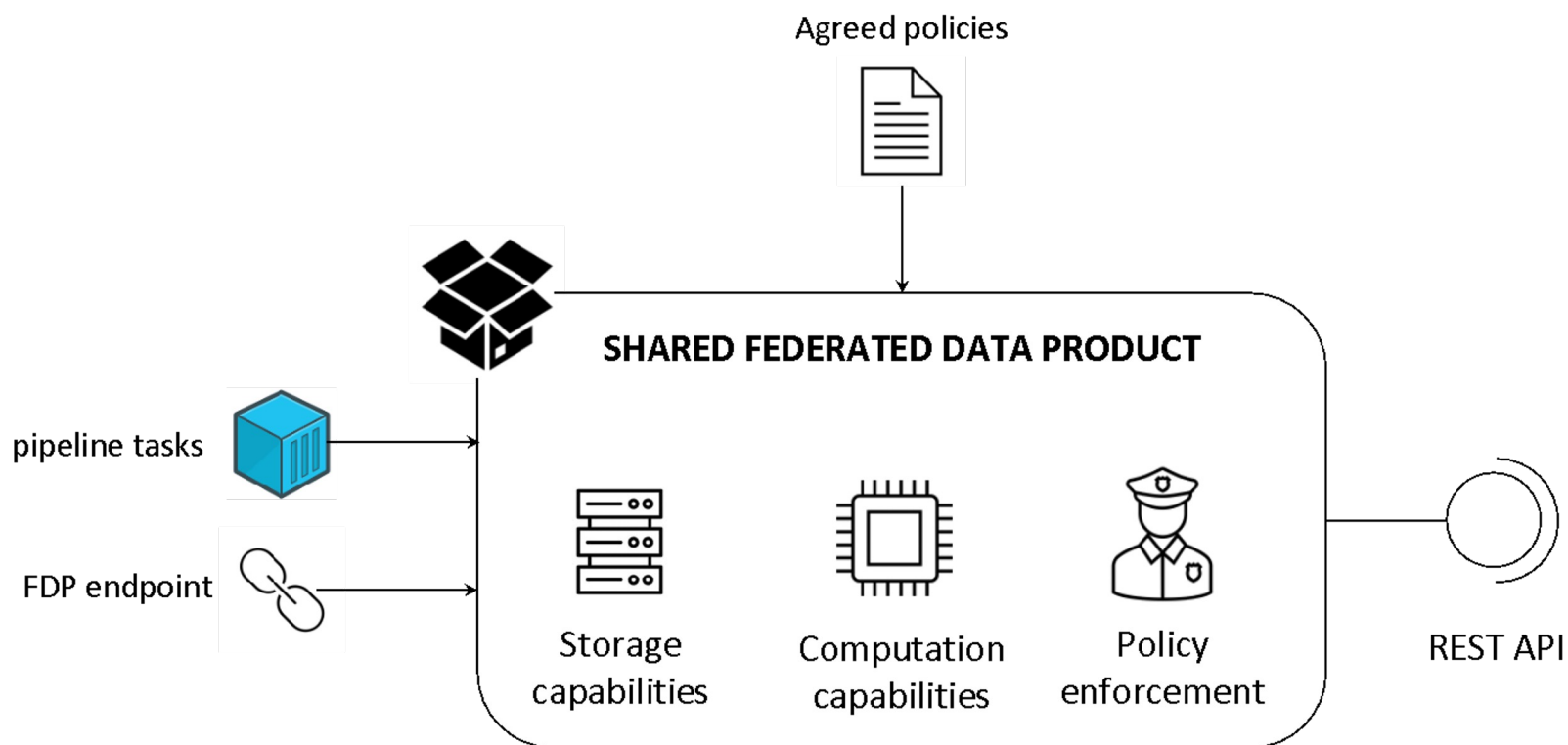
The data sharing process in TEADAL



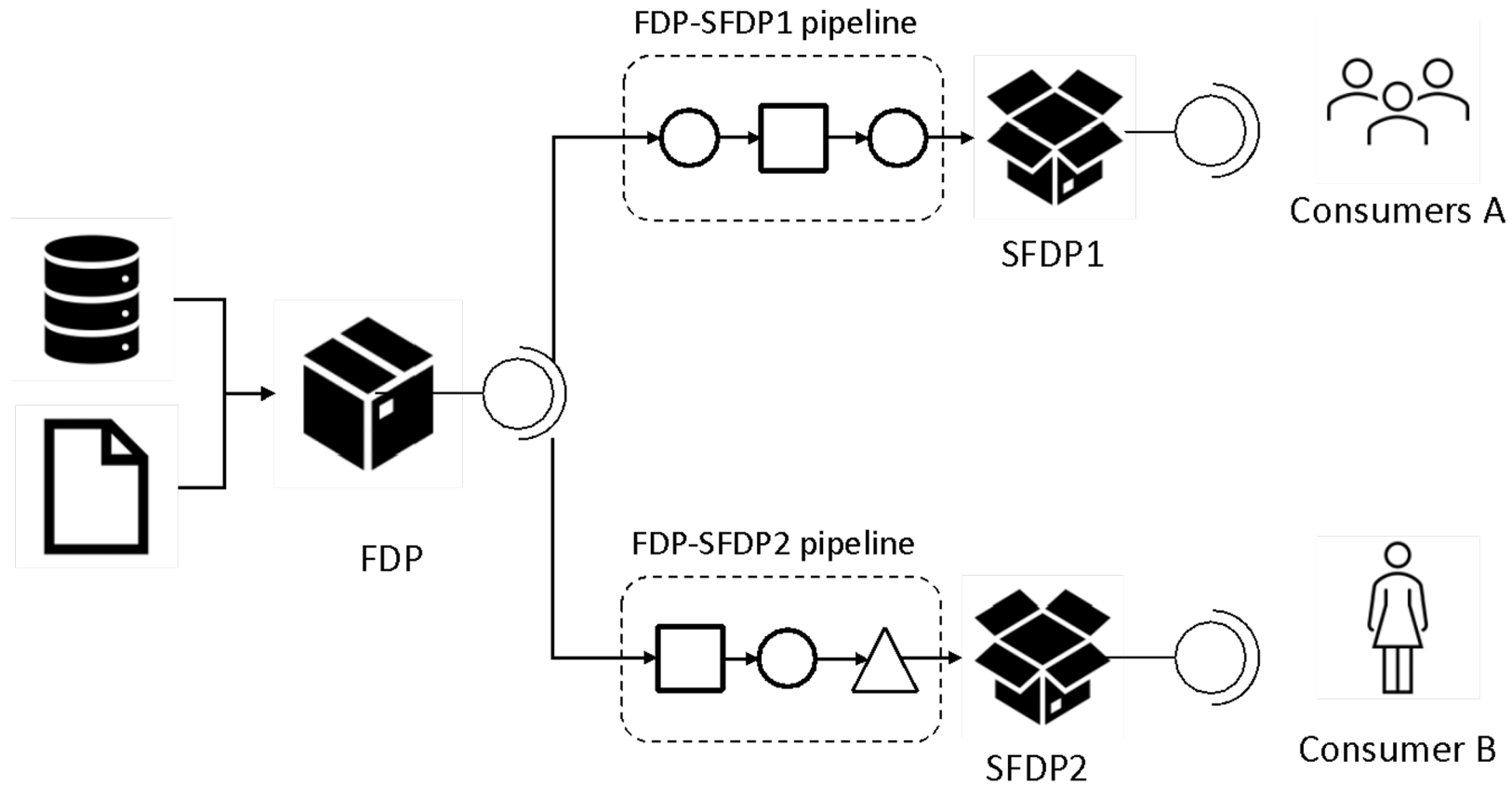
Federated Data Product



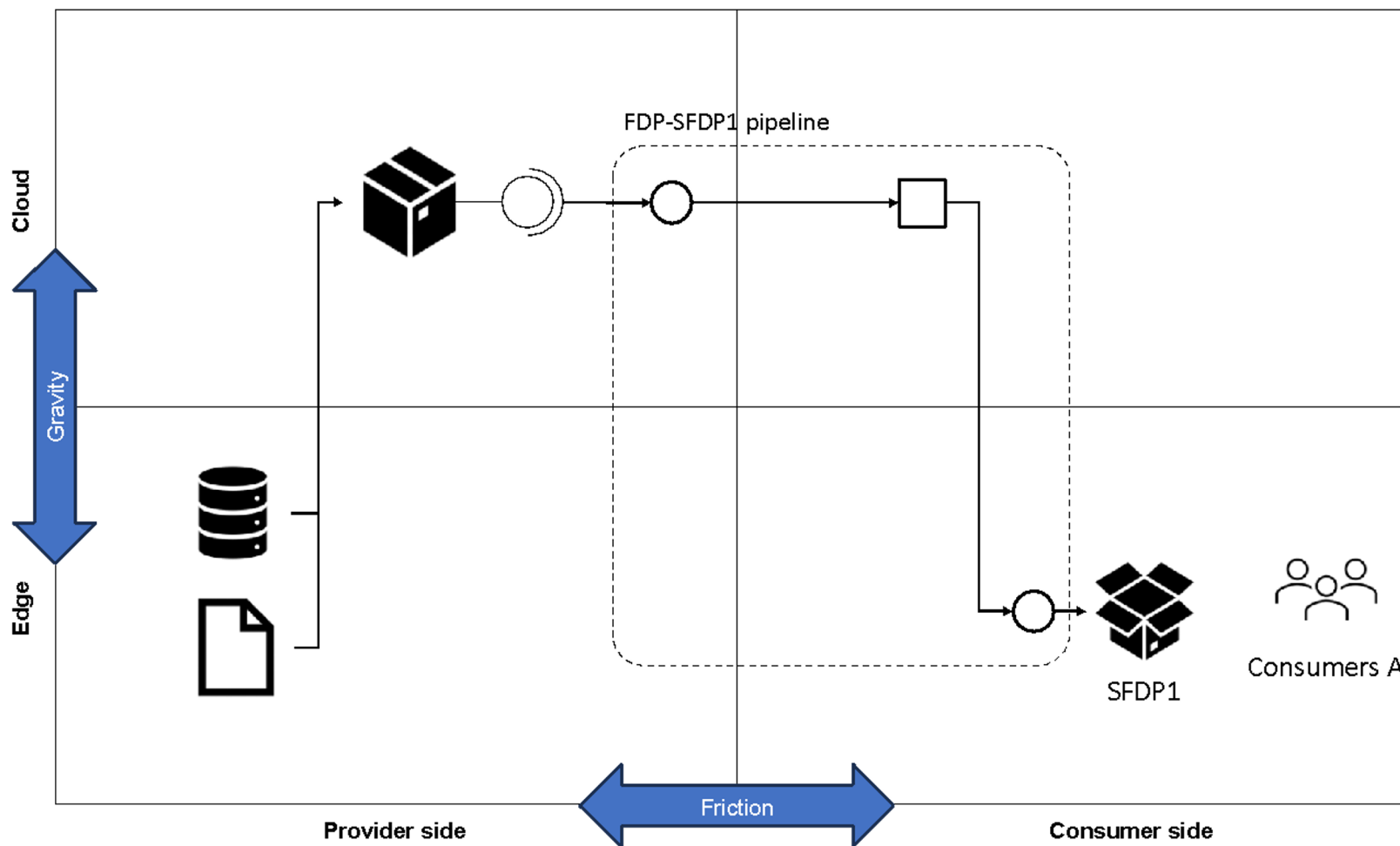
Shared Federated Data Product



The data access



Gravity and friction



Take aways

- Data mesh has been adopted in TEADAL as a pillar to define the data sharing in federated settings
- Moving outside the organization requires:
 - Mechanisms to efficiently move the data or applications
 - Ensure the trust
 - Guarantee the sovereignty
- Service orientation (which is actually already present in the data product) is the second pillar addressing the new aspects
 - In TEADAL the service orientation is made explicit

References



(Plebani et al, 2023) Pierluigi Plebani, Ronen Kat, Frank Pallas, Sebastian Werner, Giacomo Inches, Peeter Laud and Rita Santiago, Teadal: Trustworthy, Energy-Aware federated DAta Lakes along the computing continuum, CAiSE 2023 Forum

(Falconi, Plebani, 2023) Matteo Falconi, Pierluigi Plebani, Adopting Data Mesh principles to Boost Data Sharing for Clinical Trials, ICDH, 2023

(Salnitri 2023) Mattia Salnitri, Pierluigi Plebani, Towards Designing Energy-Aware Cybersecurity Policies, CAiSE Forum 2023

Meet the TEADAL Consortium



ubiwhere



Cefriel

