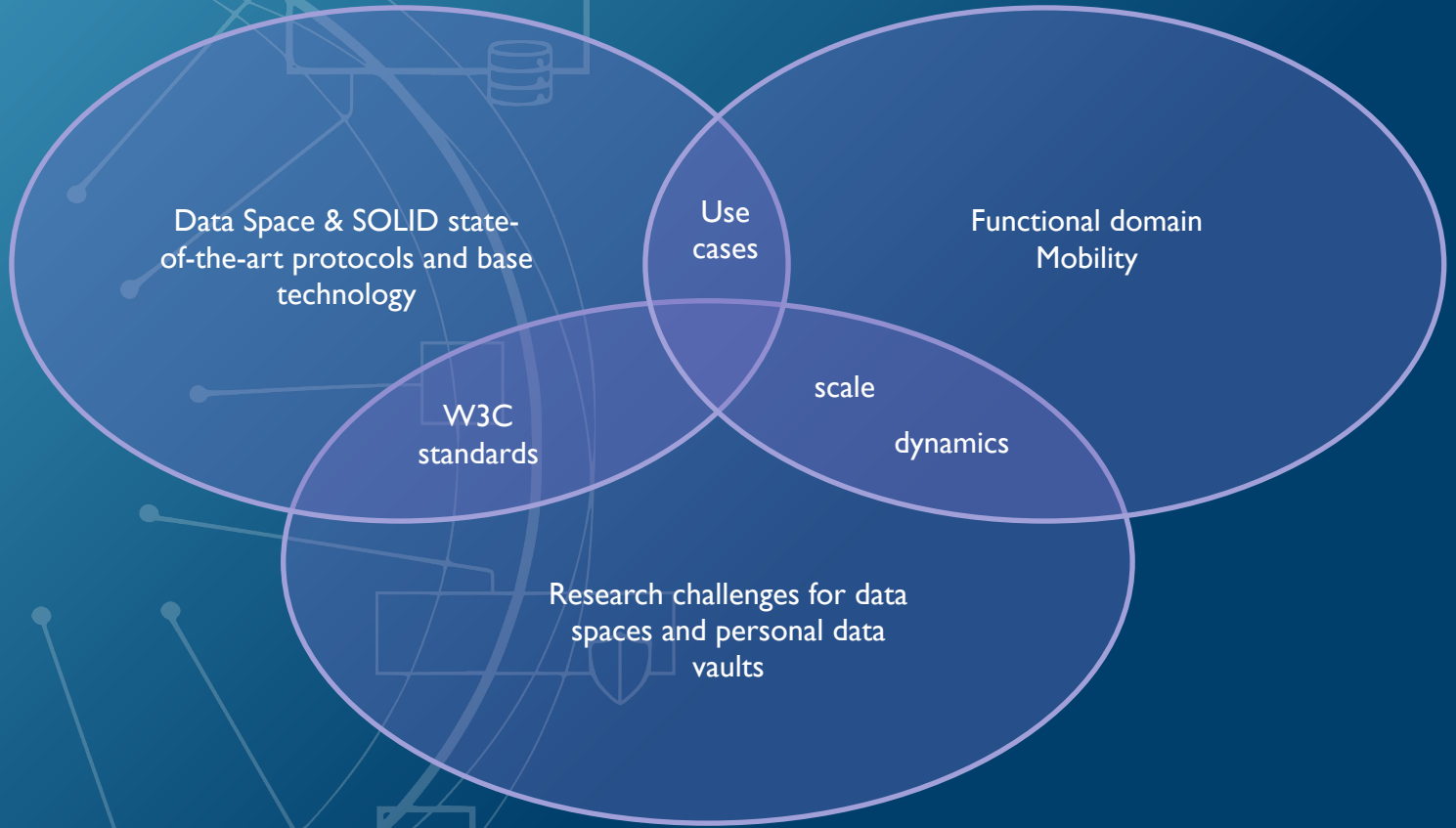
Possible *dynamics* in data sharing where Solid and Data Spaces naturally meet

Stefan Lefever – Imec AI & Algorithms

Scope


Data Space & SOLID state-of-the-art protocols and base technology
Use cases
Functional domain Mobility
W3C standards
scale
dynamics
Research challenges for data spaces and personal data vaults

Data Spaces (IDSA)

A data ecosystem, defined and governed by a sector or community, whereby **decentralized** infrastructure enables **trustworthy** and secure data sharing **capabilities** between its participants.

Note: Currently focused on B2B interactions

TRUST

- IDENTITY & CLAIMS MANAGEMENT
- ACCESS CONTROL
- SECURE & CONTROLLED EXCHANGE

INTEROP

- SEMANTICS, SCHEMAS, PROTOCOLS
- PROVENANCE & TRACEBILITY
- OPERATIONAL & LEGAL

VALUE

- DISCOVERABILITY & AVAILABILITY
- MONETIZATION
- APP & SERVICE MARKET PLACES

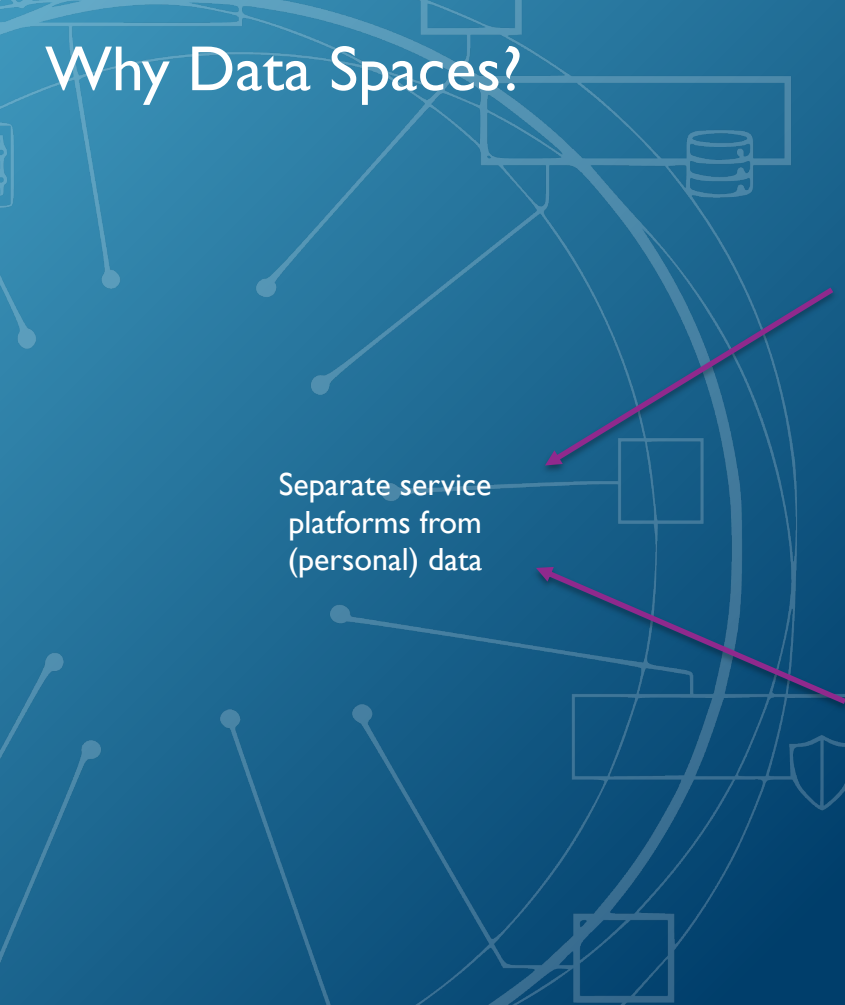
GOVERNANCE

- LEGAL FOUNDATION
- USAGE POLICY ALIGNMENT
- COMMUNITY MANAGEMENT

https://dssc.eu/space/BVE/357073747/2+Core+Concepts

imec

Why Data Spaces?



IDSA

- Create a level-playing field for data access
- Improve efficiency of data sharing and usage
- Address trust & sovereignty as main inhibitor for data sharing
- Realize an infrastructure to address the EU acts

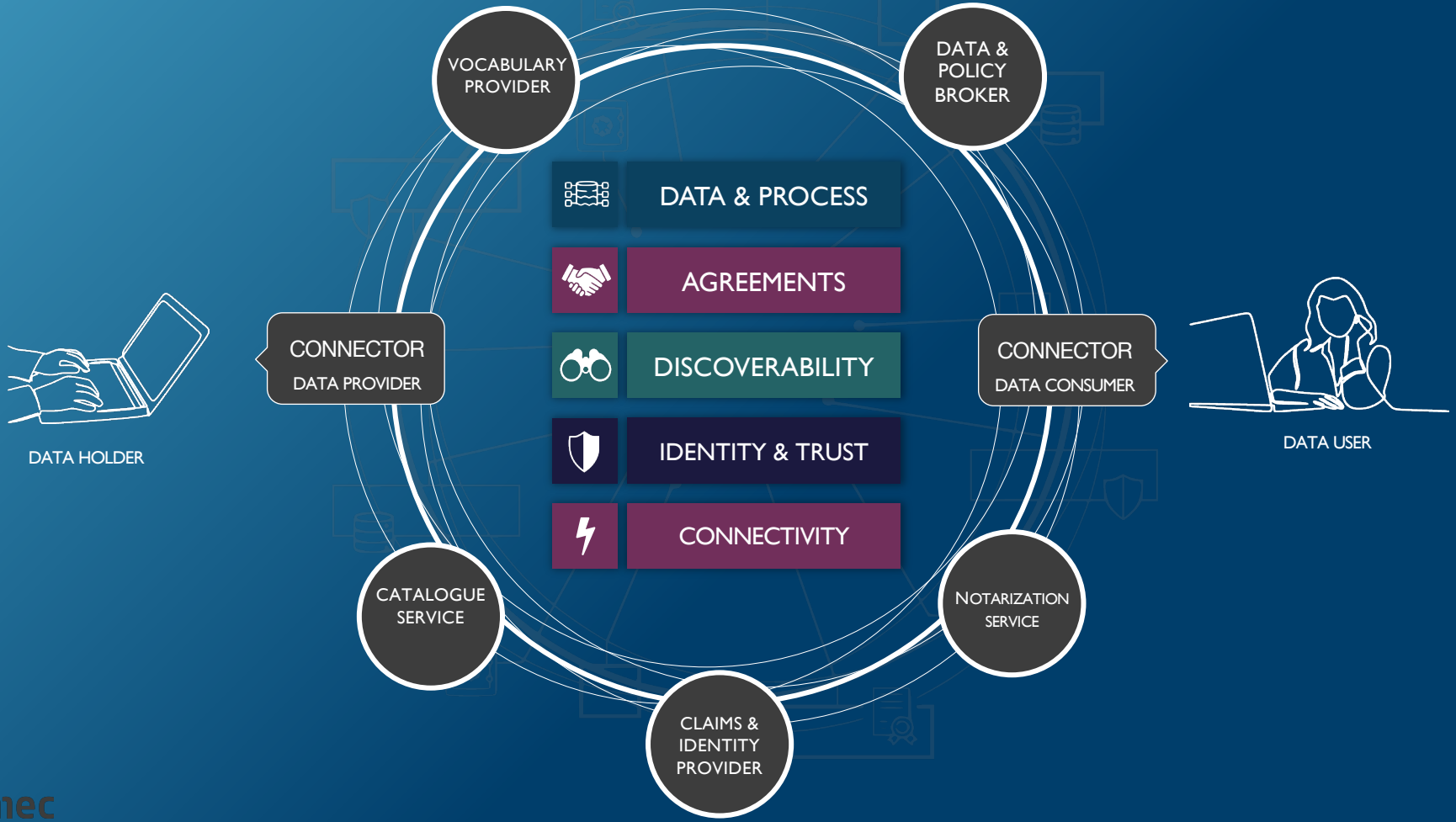
IDSA + personal data

- **IDC:** Big Data & analytics software and cloud services market reached $90 billion in 2021, will double by 2026*.
- Large part of these revenues are related to end-user query, reporting and analysis tools.
- 80% of customers indicate need for personalized experiences => access to personal data

imec

https://www.idc.com/getdoc.jsp?containerId=US50117823&pageType=PRINTFRIENDLY

Data Spaces (IDSA) : typical building blocks and capabilities

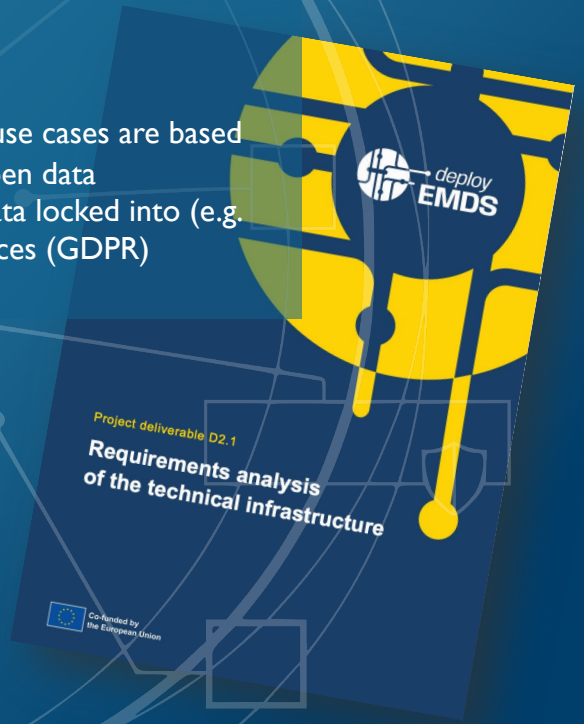


Domain : mobility

... Of Persons

Personal data value **not** fully unlocked

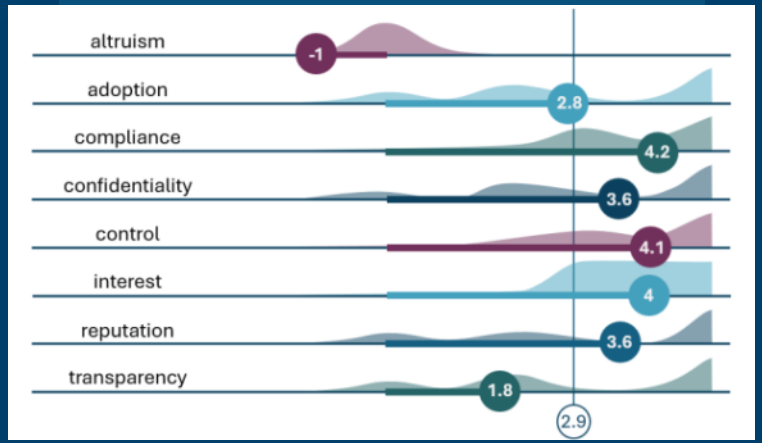
- *E.g. deployEMDS* : most use cases are based on non-personal and open data
- E.g. Personal mobility data locked into (e.g. mobility) platform services (GDPR)



... Of Goods

Data is very business sensitive
Trust is the key inhibitor but also enabler

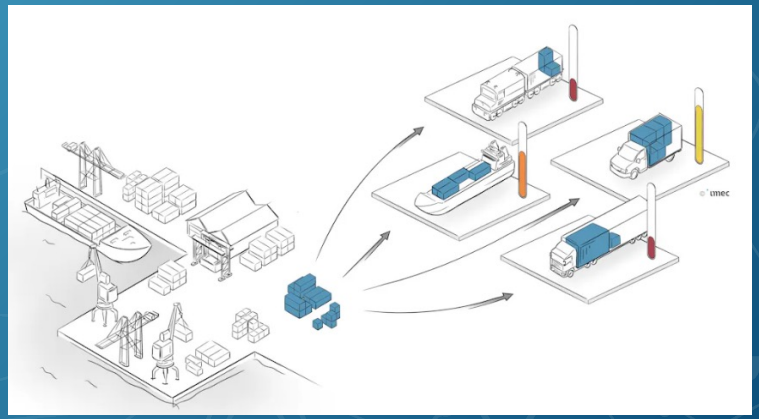
Trust is mostly driven by control, confidentiality, compliance, reputation and interest:*

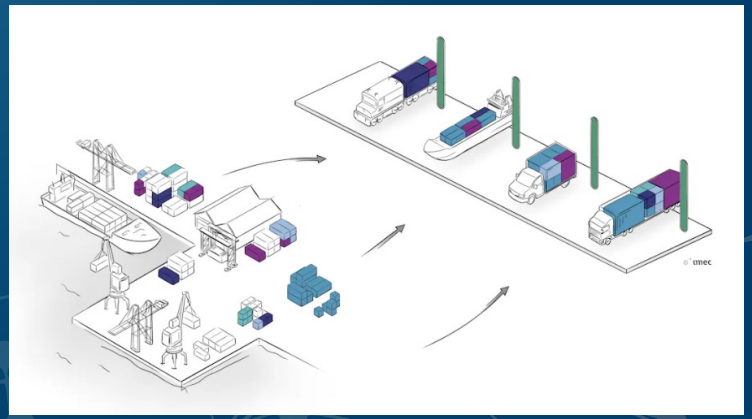


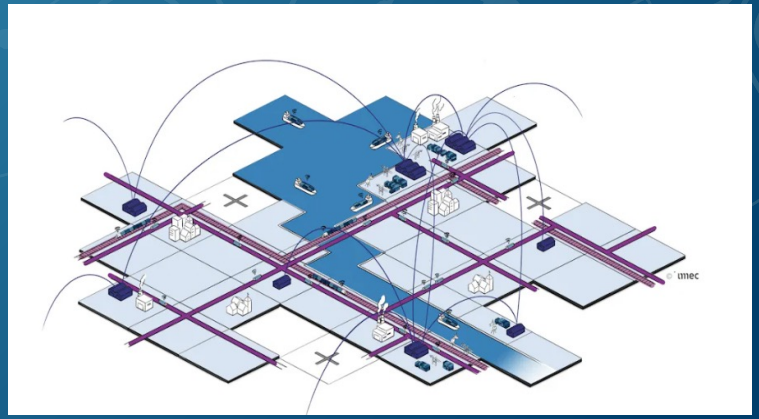
imec

** Unlocking the potential of the PI : a trust-enabling decentralized Process Sharing Connector*

Mobility of Goods : Logistics & The Physical Internet





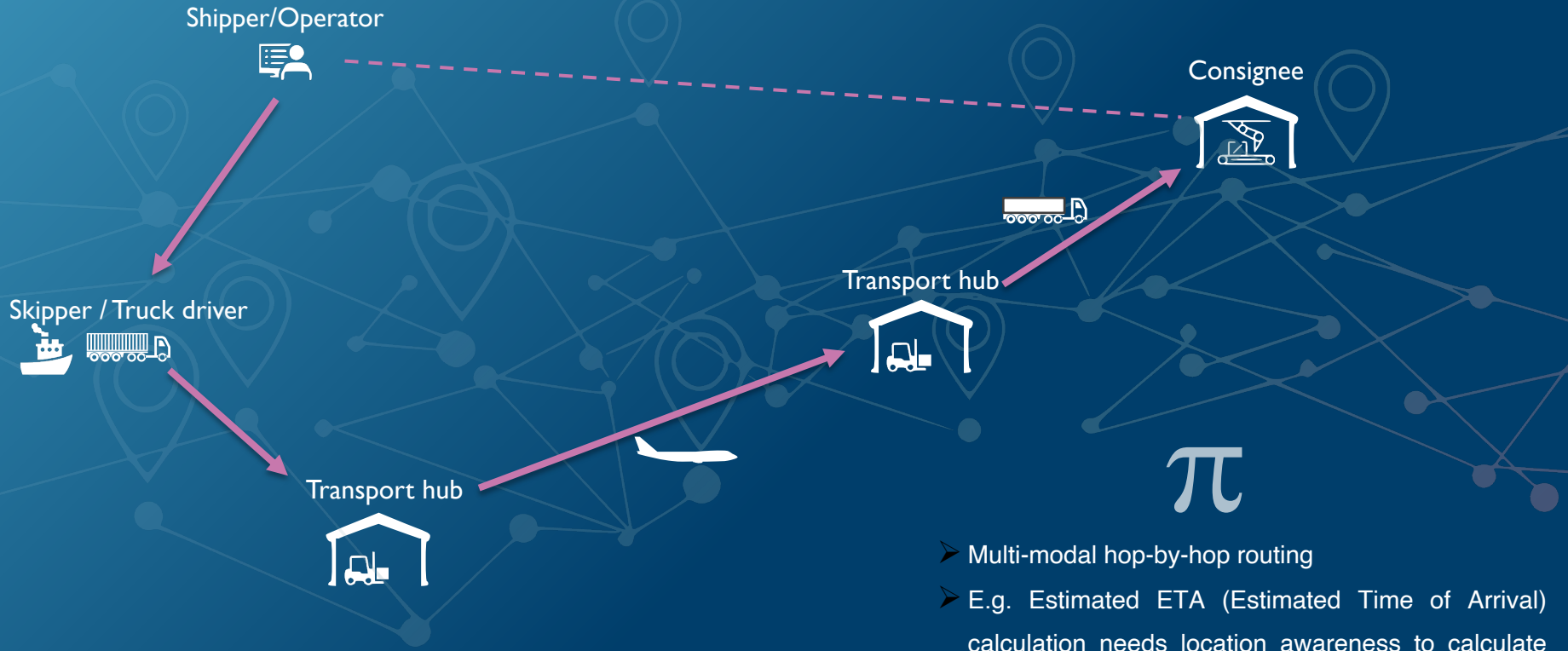


π

Analogous to the way the Digital Internet transfers data, the Physical Internet aims to make the transfer of goods more efficient, resilient and sustainable using a decentral modular open architecture. In PILL, the foundation for a trustless decentral logistics network is created. This network is the foundation of the future Physical Internet.

imec

Logistics pattern (simplified)



π

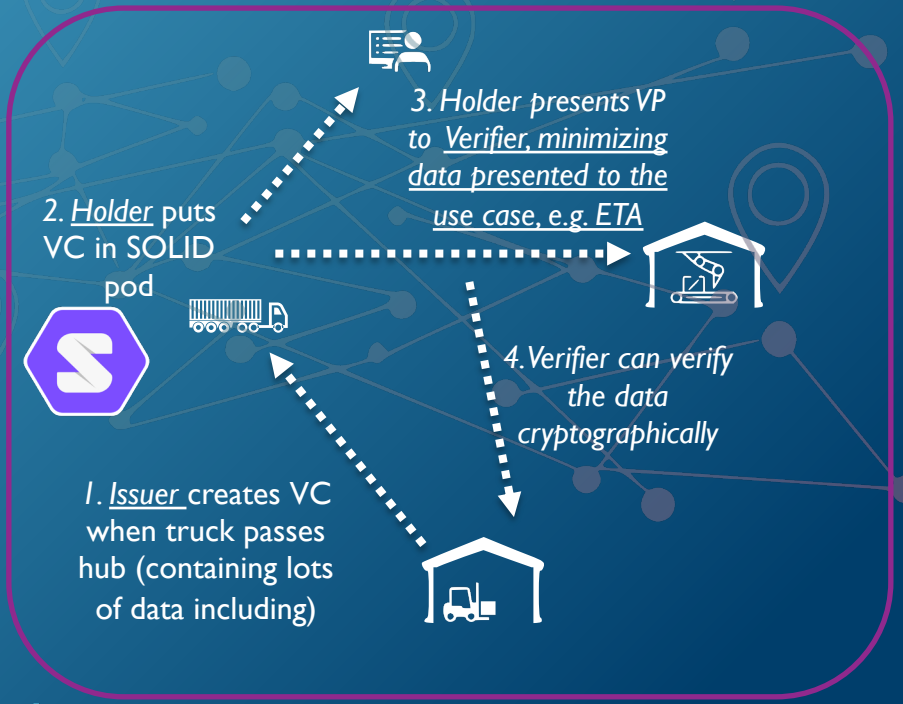
- Multi-modal hop-by-hop routing
- E.g. Estimated ETA (Estimated Time of Arrival) calculation needs location awareness to calculate next-hop routing

imec

Scenario: trusted & verifiable reporting using Verifiable Credentials

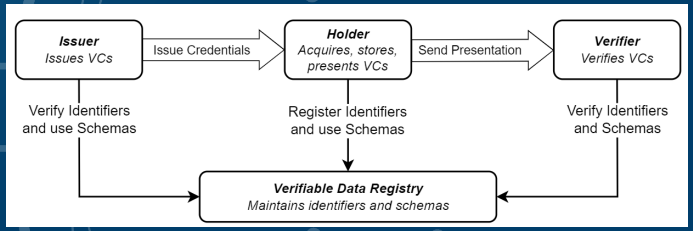
Question : Shipper/Consignee wants to verifiably track ETA by being informed on hub transfer of truck

Problem : location of transport is also home residence location



Data Spaces

- Rely heavily on W3C Verifiable Credentials and Presentations
  - For Authentication (*remark both SOLID and data spaces are built on decentralized IDs*)
  - For Verifiable Data Presentation and Minimization
- Using agreed upon data schema's and verifiable data registries



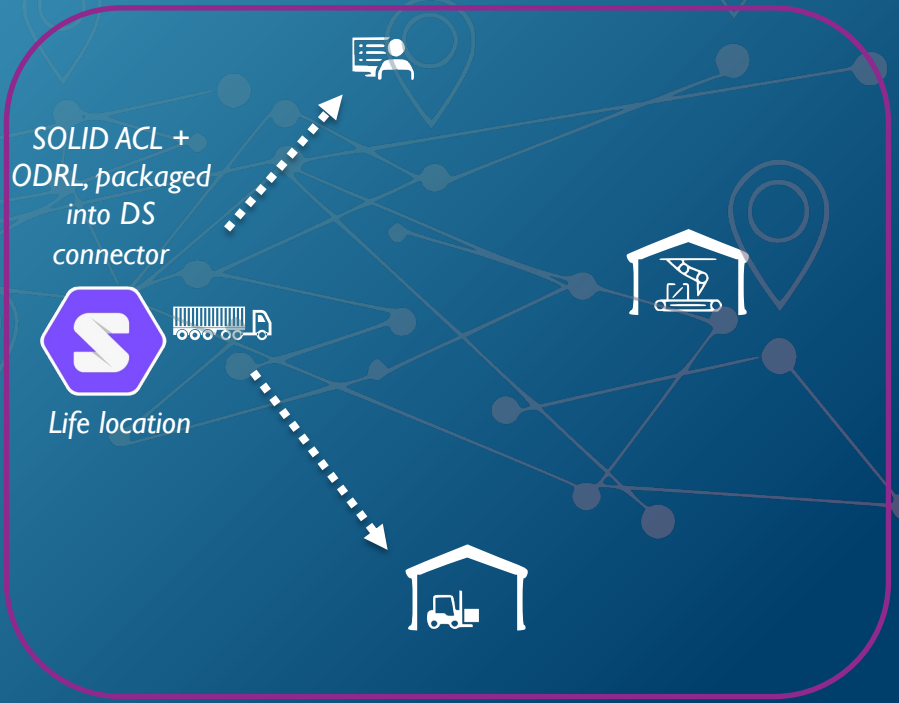
Solid

- SOLID PODS could be used to store Credentials
- **Derived VC** can be used to translate static data in the pod (e.g. birthdate) into dynamic presentations (e.g. older than 18) using ZKP algorithms.
- Model access control rules to express which attributes can be used in derived credentials

imec

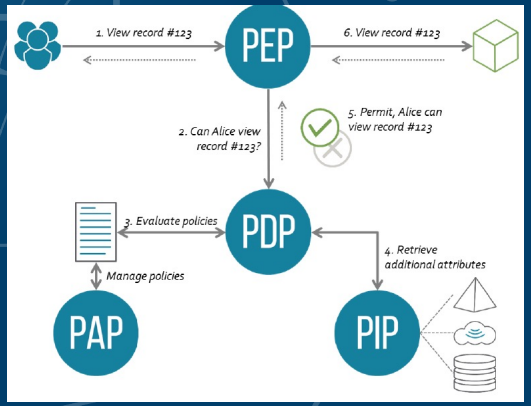
Scenario: usage control and policies

Question : How can access to data be automated without interactions of humans, but taking the current context in mind (e.g. working hours) ?



Data Spaces

- Contract Negotiation Protocol relies heavily on policy definition, management and enforcement, using ODRL as proposed language.



Solid

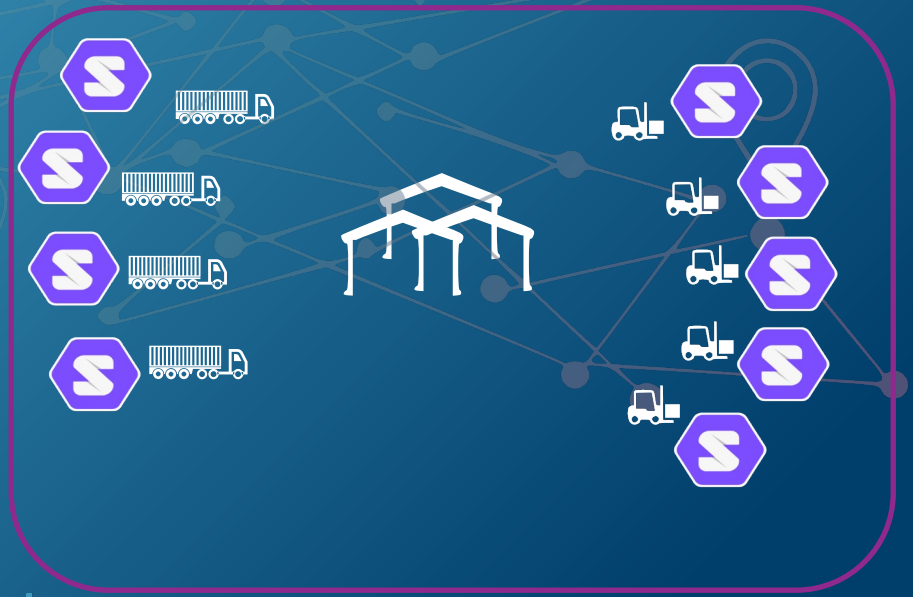
- ODRL + DPV extensions to ACL
- Extensions to support dynamic policies in ODRL
- Consent in GDPR : PDI (Personal Data Intermediaries)
- Integration with a DS Connector
- Extension to the IDSA Data Space Protocol ?

imec

Scenario: decentralized / federated querying

Question : *How can a hub have access to data from all trucks and containers when truck enters ?*

Question : *How can a truck detect hub services when entering ?*



Data Spaces

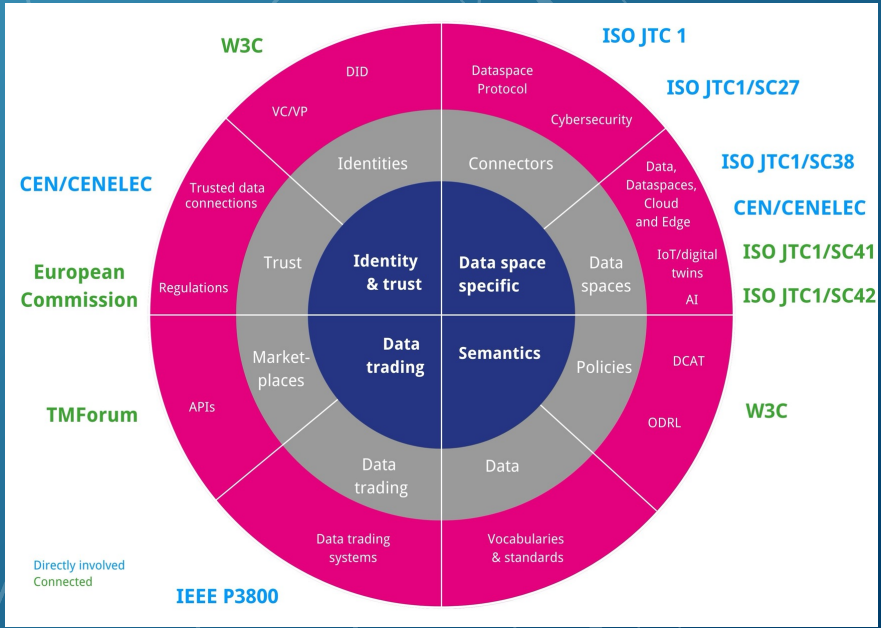
- Central or federated catalogues & metadata brokers
- But these do not scale very well in very sovereign and dynamic scenario's
- Hubs can local connector registries (e.g. to offer access to local SOLID pods)

Solid

- Using the **power of Linked Data** (e.g. through Linked Traversal Query Processing), the locally registered pods can be queried in a federated way.
- Through the pod of the truck, links can be found to the PODs of the containers on the truck.
- And all of this without central catalog needed which copies the data continuously (all data is up-to-date in each POD)

imec

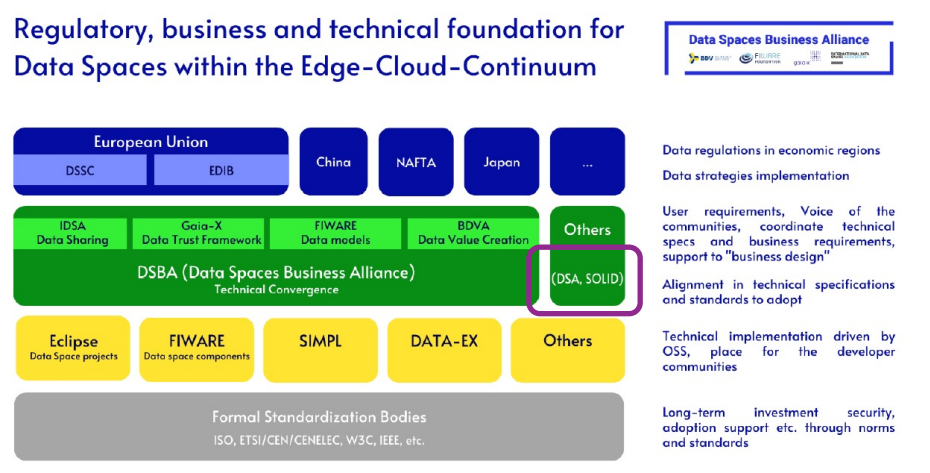
Conclusions :



IDSA standardization radar

- SOLID and WEB technology in general **align very well** with (IDSA) choices in data space technology.
- **Value of fusing personal** data and business data is increasing
- Implementations of regulations like GDPR, Data Act, Data Governance Act, AI act,... becomes a challenge, but technology can assist in addressing that.
- Some synergizing issues (techno vs techno, DS techno vs SOLID techno) are not solved yet and need some research
- Aligning the Data Product thinking of DS with SOLID can help to convergence the principles.
- Some aspects of SOLID and DS require more dynamics to scale.

Summary : momentum



Existing Software Landscape

Likely incomplete, no specific judgement or priority in the order of representation

Gaia-X

- **Gaia-Lab** Registry, Credential issuers, Policy based implementation of conformity, Catalogue synchronization
- **Eclipse XFSC** SSI, Federated Catalogues, Notarization Services, Wallets, Workflow
- **Prometheus-X** Personal Consent

IDSA

- **Connector Report** 14 OSS based IDSA Connector implementations (28 total)
- **Basecamp** e2e implementation of RAM 4.0 (incl. DAPS, Clearing House, Metadata Broker...)

FIWARE

- FIWARE Connector
- FIWARE OSS Marketplace
- NGSI Context Brokers
- DOME Marketplace
- Smart Data Models

Other

- **Eclipse Data Space Components**
- **Pontus-X** Compute2Data, Ocean Protocol Provider, Ocean Aquarius
- **iShare** Identity, Legal Framework
- **Solid** Personal Data spaces
- **Tractus-X** Semantic Data Hub, Digital Twin registries...
- Various **Wallets**

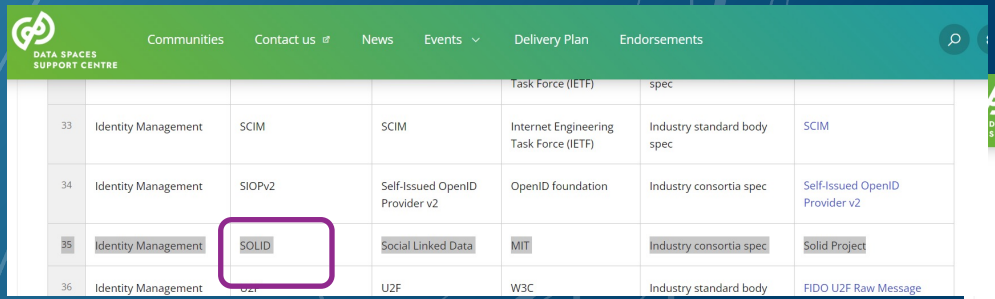
Commercial

- **Dawex** Data Contracting, Data Marketplace
- **Eviden** Financial Clearinghouse

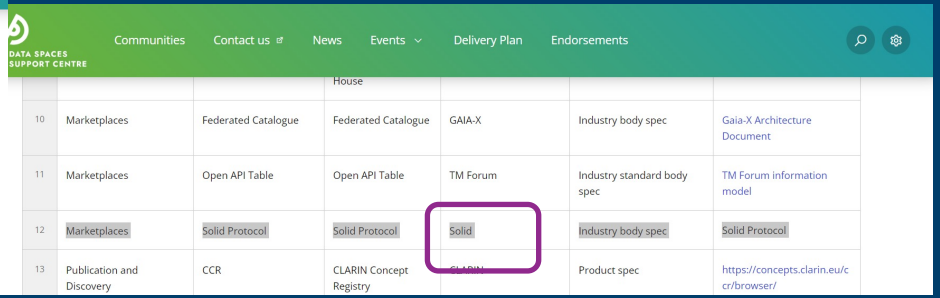
EU Building Blocks

- **eDelivery**
- **eID**
- **eSignature**
- **Context Broker**
- **EBSI**
- **...**

EVIDEN



				Task Force (IETF)	spec	
33	Identity Management	SCIM	SCIM	Internet Engineering Task Force (IETF)	Industry standard body spec	SCIM
34	Identity Management	SIOPv2	Self-Issued OpenID Provider v2	OpenID foundation	Industry consortia spec	Self-Issued OpenID Provider v2
35	Identity Management	SOLID	Social Linked Data	MIT	Industry consortia spec	Solid Project
36	Identity Management	U2F	U2F	W3C	Industry standard body	FIDO U2F Raw Message



				House		
10	Marketplaces	Federated Catalogue	Federated Catalogue	GAIA-X	Industry body spec	Gaia-X Architecture Document
11	Marketplaces	Open API Table	Open API Table	TM Forum	Industry standard body spec	TM Forum information model
12	Marketplaces	Solid Protocol	Solid Protocol	Solid	Industry body spec	Solid Protocol
13	Publication and Discovery	CCR	CLARIN Concept Registry	CLARIN	Product spec	https://concepts.clarin.eu/ccr/browser/

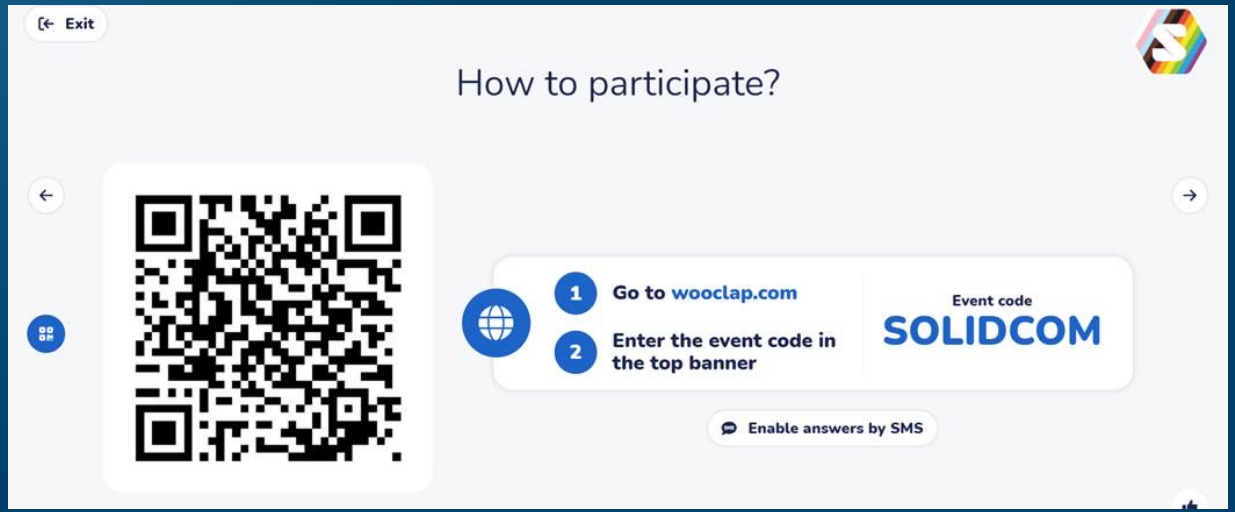
imec

Summary : panel statements

1. Without **personal** data, the (mobility) data space will never **scale** and always be stuck in ‘closed ecosystems’ with only limited use cases.

2. The current specifications that are building up the Solid & IDSA ecosystems have the **potential to synergize**, but need to address real shortcomings, especially in the field of **more dynamics and automation** in sovereign data access.

3. In the current “**shift to the edge**” paradigm, private data vault technology needs to scale to the edge and integrate seamlessly with time series data and edge-based (or distributed) algorithmic power.



imec

Exit

How to participate?

1 Go to wooclap.com

2 Enter the event code in the top banner

Event code

SOLIDCOM

Enable answers by SMS

embracing a better life