

Nov 25, 2025 | 10:00 - 11:30 CET

Data governance, Data quality, Data spaces maturity and the path to Standardization



FAME



INTERNATIONAL DATA
SPACES ASSOCIATION

Welcome

Why we have this webinar



Context

- Europe has set out clear rules for a fair and open data economy. The **Data Act**, the **Data Governance Act** and the **AI Act** define who can access data, under what conditions and for what purposes
- **Standardisation** is needed to make these rules work in practice
- The **EUDATA+** Cluster unites seven projects on data monetisation and exchange, fostered by the European Commission's collaboration programmes like Horizon Europe



Motivation

Objectives of the workshop:

- **create awareness** on the **European standardisation** activities around **Data** and **Data spaces**
- **identify synergies** among the projects and **CEN/CENELEC** working groups working on these topics to spark collaboration and contributions

Who

Standardisation experts



Clara Pezuela Robles



Ismael Caballero Muñoz-Reja



Silvia Castellvi



Who

Experts on Data governance, Data quality and Data spaces

Projects



Speakers



Ernesto Troiano



Jordi Arjona Aroca



Yury Glikman



Agenda

Insights from CEN/CENELEC JTC 25 30'

- Introduction Silvia Castellvi, IDSA
- WG 2 Dataspace - Maturity assessment of Common European Data Spaces Clara Pezuela Robles, ITI
- WG 3 Data management and governance - Quality assessment of internal data governance processes Ismael Caballero Muñoz-Reja, UCLM

Learnings from projects 30'

- FAME Ernesto Troiano , GFT
- DATAMITE Jordi Arjona, ITI
- PISTIS Yury Glikman, FHG FOKUS

Open discussion 20'

- **moderated by** Silvia Castellvi, IDSA

Next steps 10'

- Silvia Castellvi, Ilknur Chulani, IDSA

Agenda

Insights from CEN/CENELEC JTC 25 30'

- [Introduction Silvia Castellvi, IDSA](#)
- WG 2 Dataspaces - Maturity assessment of Common European Data Spaces Clara Pezuela Robles, ITI
- WG 3 Data management and governance - Quality assessment of internal data governance processes Ismael Caballero Muñoz-Reja, UCLM

Learnings from projects

Open discussion

Next steps'

Agenda

Insights from CEN/CENELEC JTC 25 30'

- Introduction Silvia Castellvi, IDSA
- [WG 2 Dataspaces - Maturity assessment of Common European Data Spaces](#) Clara Pezuela Robles, ITI
- WG 3 Data management and governance - Quality assessment of internal data governance processes Ismael Caballero Muñoz-Reja, UCLM

Learnings from projects

Open discussion

Next steps'

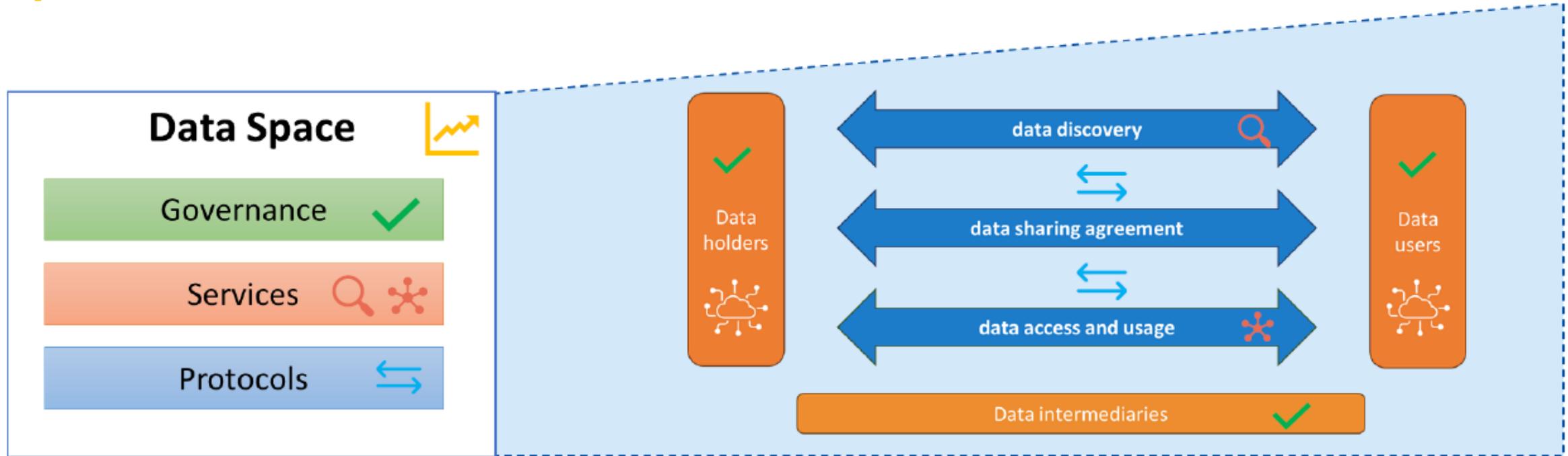
Standardization on Maturity Assessment of Data Spaces

Clara Pezuela

Expert in JTC25 – Project Lead for TS on Maturity Model

Standardisation request

European Trusted Data Framework



 Trusted Data Transaction standard

 Data catalogue implementation framework

 Data governance standard for data space participants

 Semantic assets implementation framework

 Maturity model for Common European Data Spaces

EU Trusted Data Framework standardization request (2025-07-01)

EU Data Act | Article 33, Essential requirements regarding interoperability of data, of data sharing mechanisms and services, as well as of common European data spaces

*Participants in data spaces that offer data or data services to other participants shall comply with the following essential requirements to facilitate the interoperability of data, of data sharing mechanisms and services, as well as of common European data spaces which are purpose- or sector-specific or cross-sectoral **interoperable frameworks for common standards and practices** to share or jointly process data for, inter alia, the development of new products and services, scientific research or civil society initiatives(...)*
*The Commission shall, pursuant to Article 10 of Regulation (EU) No 1025/2012, **request one or more European standardisation organisations to draft harmonised standards** that satisfy the essential requirements laid down in paragraph 1 of this Article*

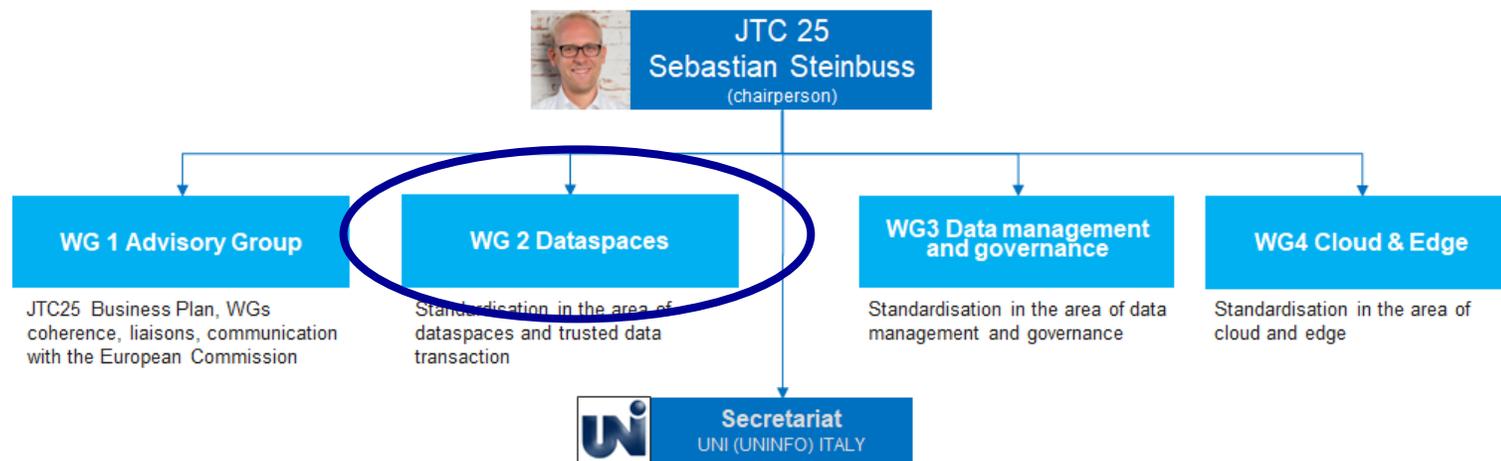
Ref.	Deliverable Title	Deadline	
1	Harmonised standards on Trusted Data Transactions — Part 1: Terminology, concepts and mechanisms	2026-06-01	CENELEC
2	Harmonised standards on Trusted Data Transactions — Part 2: Trustworthiness requirements	2026-11-01	CENELEC
3	Harmonised standards on Trusted Data Transactions — Part 3: Interoperability requirements	2027-05-01	CENELEC
4	Technical specification(s) on a data catalogue implementation framework	2026-03-01	ETSI
5	Technical specification(s) on an implementation framework for semantic assets	2026-09-01	ETSI
6	European standard on a quality framework for internal data governance	2027-03-01	CENELEC
7	Technical specification(s) on a maturity model for Common European Data Spaces	2026-09-01	CENELEC



CEN/CENELEC JTC 25



- Launched on the 20th September 2024 in Brussels
- Scope on data management, dataspaces, cloud and edge
- More than 100 experts across EU
- Growing liaisons with projects and associations



Measuring the maturity of a data space is not a trivial task and required certain guidance



Which is the context of the Technical Specification (TS)?

- Data Spaces are key enablers for data sharing according to EU regulation (Data Act, Data Governance Act)
- European Trusted Data Framework standardization request to support the Article 33 of Data Act
- Includes a deliverable on Technical Specification(s) on a maturity model for Common European Data Spaces
- The Data Spaces Support Centre (DSSC) has produced a Maturity Model which has taken as baseline for the TS



Work in progress...



Which is the scope of the TS?

- Define a maturity model concept, structure, methodology and measurable criteria
- And related requirements and guidance for the assessment of data space maturity

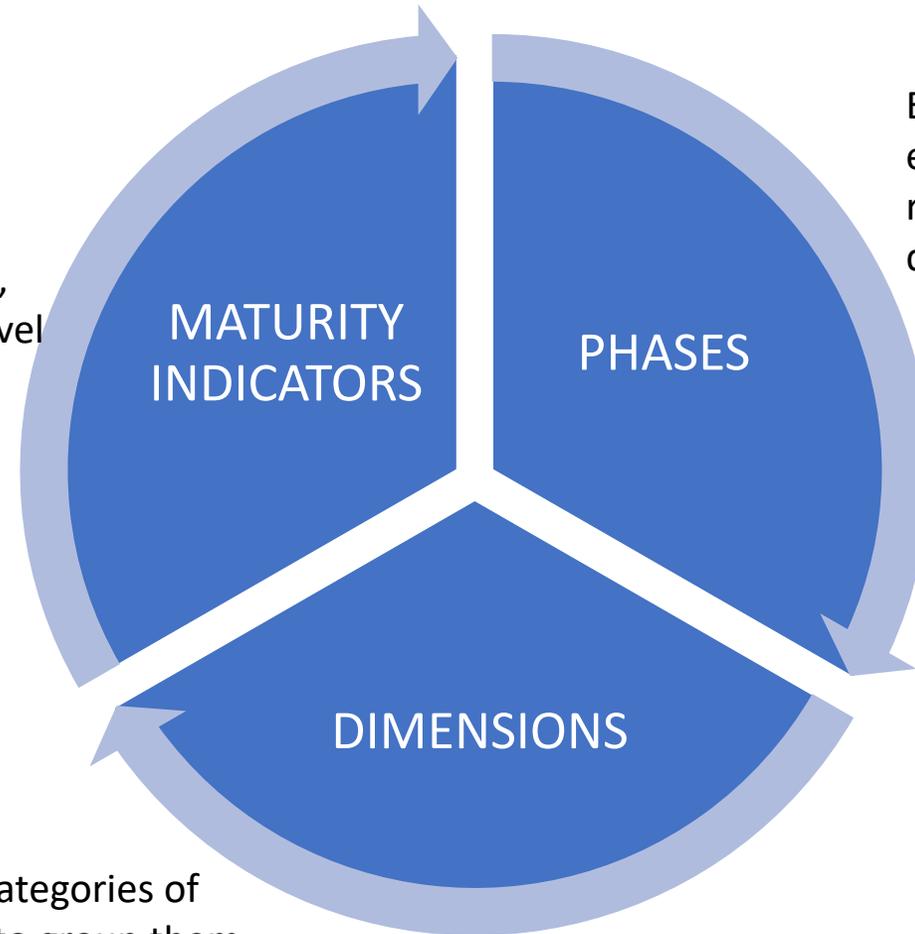


Which are the intended use of the TS?

- As a Data Space (or any other data sharing initiative), you can:
 - self-assess the evolution of your data space at any time
 - compare your data space with the others (if self-assessments are public)
 - identify your data space strengths and areas for improvement to evolve
- As candidate to join a Data Space, you can decide your engagement based on the published self-assessment
- As Policy Maker, you can get an overview of all data spaces maturity and take actions based on analysis
- As a Certification Provider, you can use the model for creating your certification offering



Elements in the maturity model



Quantitative or qualitative metrics to monitor, measure, and evaluate the maturity level of a data space

Each of the steps in the evolution of a data space, representing a specific level of development and capability

- Exploratory
- Preparatory
- Implementation
- Operation
- Scaling

- Business
- Governance
- Legal
- Interoperability
- Control over data & Trust
- Value Creation

Different categories of indicators to group them



Reporting structure with additional information



Template for self-assessment



Data Space Maturity Indicators (DSMI)

Indicator description

- Which aspect of the data space do we want to assess?

Metrics (driven by capabilities)

- What do we want to measure? Which are the included capabilities? How do we ask it?

Measurable values

- Which are the possible values to measure?

Scoring per phase

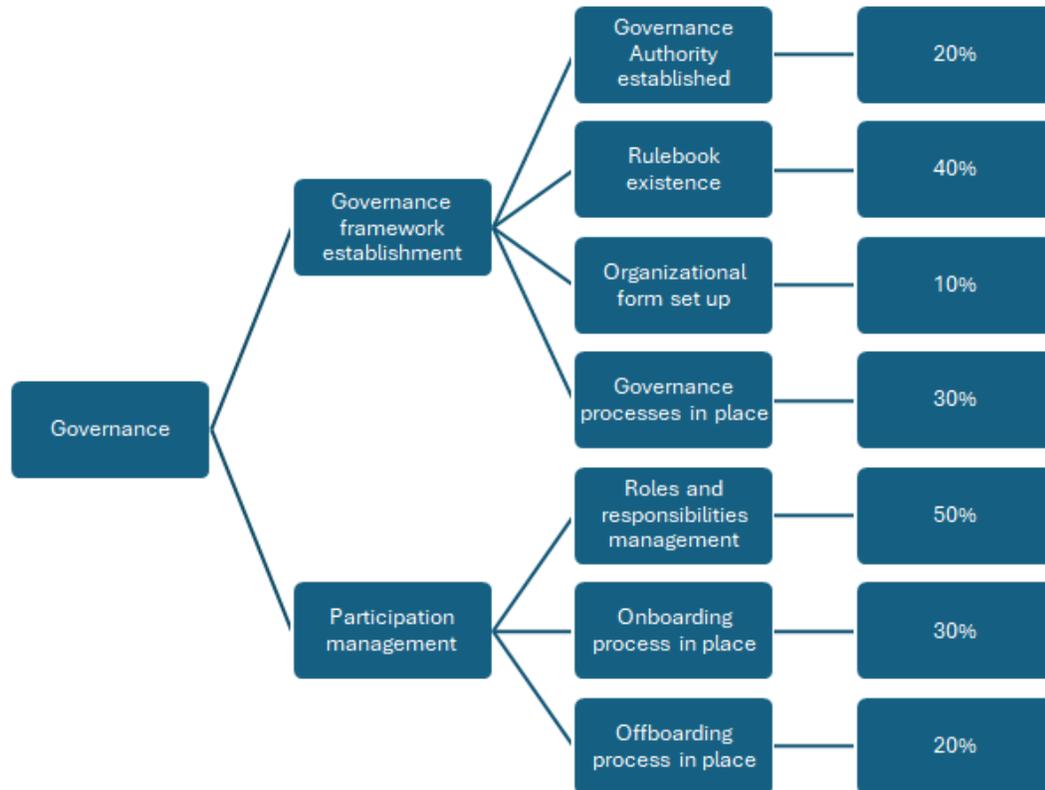
- How do we want to measure it?

Weight of relevance

- How relevant is every maturity indicator within the overall dimension? (the total must be 100%)



Example of Governance maturity assessment



Only Conceptual	Already Drafted	Established but still under validation	In continuous operation	Even enable to scale
Not exists	Exists with basic tailoring	Comprehensive, tailored, and covers all key topics	Used in daily operation at least in human-readable format	Fully operational, digital integrated, machine-readable format
Not defined yet	Drafted	Established	Enforced	Size-adapted structure
No formal processes yet	Drafted processes	Processes defined and communicated	In enforcement, regularly updated	Size-adapted processes
Not defined yet	Drafted	Established	Continuously running	Size-adapted
Not defined yet	Drafted	Established	Continuously running	Size-adapted
Not defined yet	Drafted	Established	Continuously running	Size-adapted



Template for self-assessment (example)

Maturity Assessment Self-Assessment template

* Indica que la pregunta es obligatoria

Governance dimension

What is the Governance Authority (GA) level of establishment in your data space? *

The GA is the body of a particular data space, responsible for developing, maintaining, operating and enforcing the governance framework. The expected capabilities are: Governance framework creation; governance model definition; ensure transparent decision making process; internal regulations in place.

- Only conceptual
- Already drafted
- Established but still under validation
- In continuous operation
- Even enable to scale

Sin título

[Atrás](#) [Enviar](#) [Borrar formulario](#)

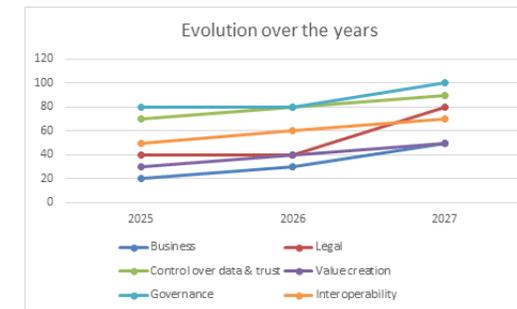
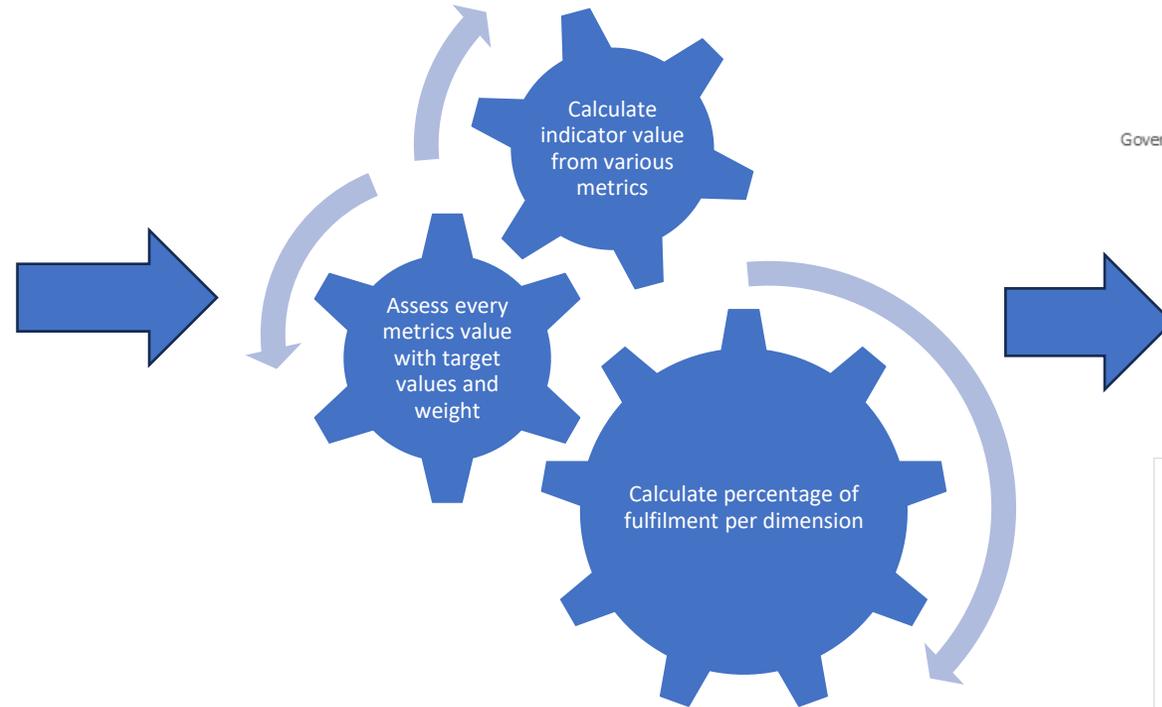
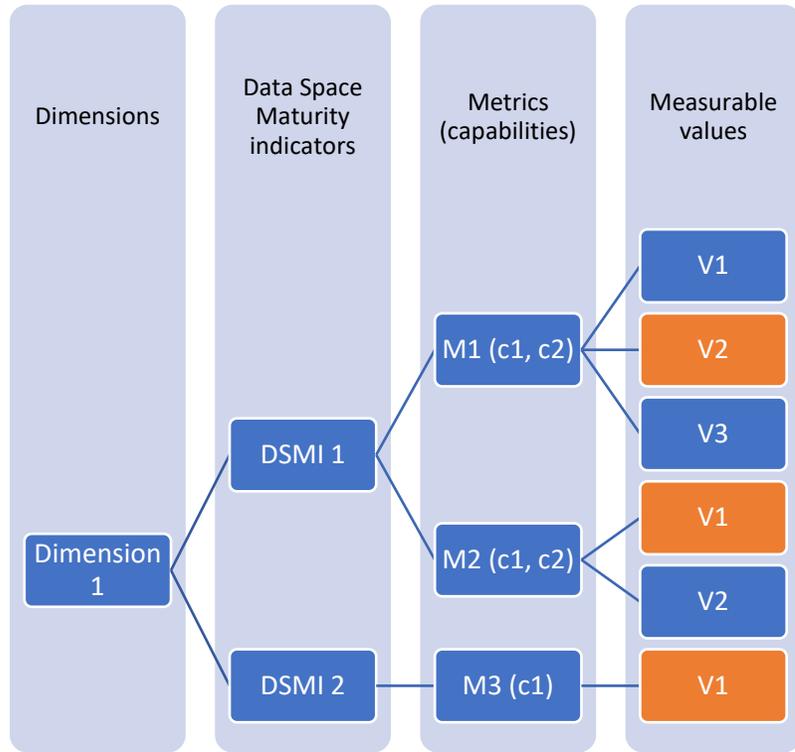
Este contenido no ha sido creado ni aprobado por Google. - [Contactar con el propietario del formulario](#) - [Términos del Servicio](#) - [Política de Privacidad](#)

¿Parece sospechoso este formulario? [Informe](#)

Google Formularios



Maturity assessment method (in progress)



Questionnaire

Analysis

Visualization

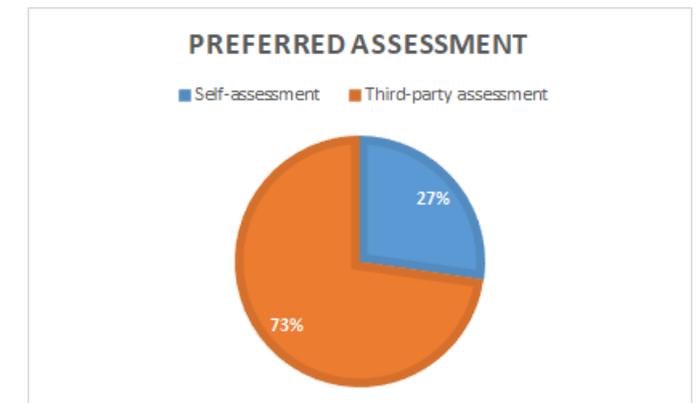
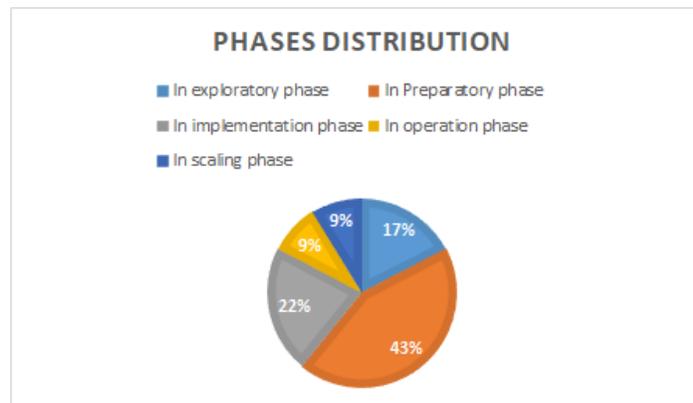
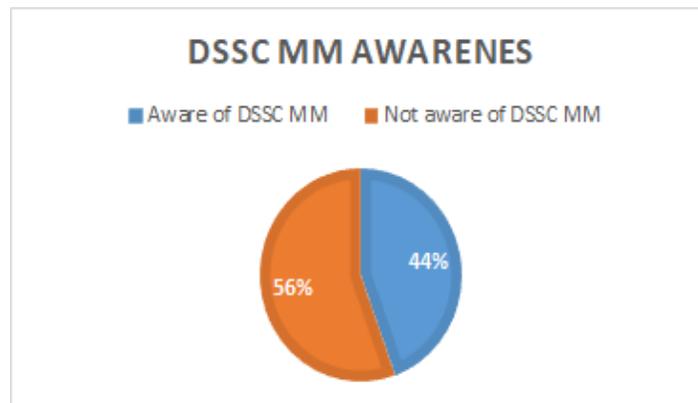
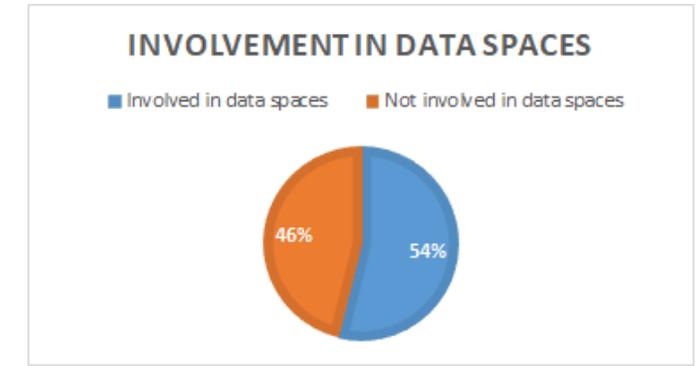


No matter your phase,
matter your continuous evolution

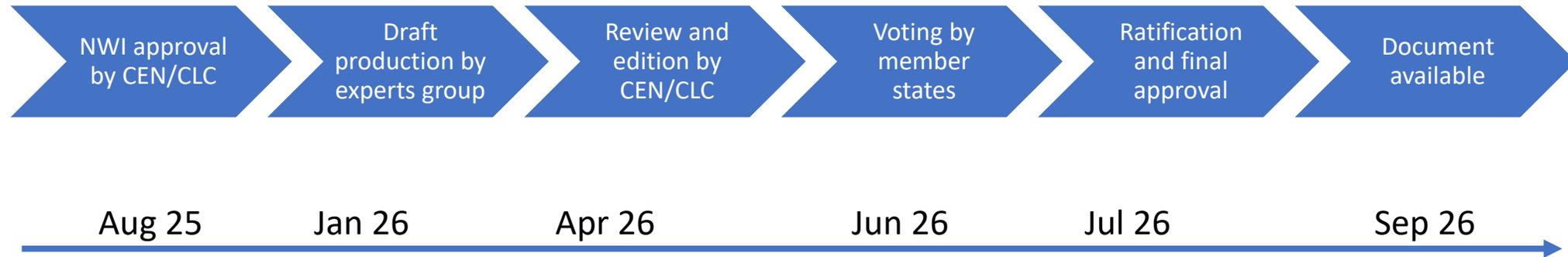


Workshops with community

- At EBDVF 2025 in Copenhagen and GAIA-X Summit in Porto
- 50 participants in total
- The dimension most relevant for measuring the maturity is Value Creation, followed by Governance
- Provided insights about what to measure: ROI, compliance, standards, use cases, transparency....

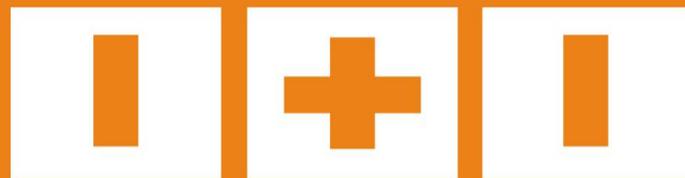


Timeline for TS production



STAY TUNED FOR NEXT STEPS!!





ITI INVESTIGATE
TO INNOVATE

Camino de Vera, s/n CPI-UPV
Ed/ Bldg. 8G Acc. B - Nivel 4 / 4th Floor
46022 Valencia - España/Spain
Telf.: +34 96 387 70 69

www.iti.es



Agenda

Insights from CEN/CENELEC JTC 25 30'

- Introduction Silvia Castellvi, IDSA
- WG 2 Dataspaces - Maturity assessment of Common European Data Spaces Clara Pezuela Robles, ITI
- WG 3 Data management and governance - Quality assessment of internal data governance processes Ismael Caballero Muñoz-Reja, UCLM

Learnings from projects

Open discussion

Next steps'

Agenda

Insights from CEN/CENELEC JTC 25

Learnings from projects 30'

- FAME Ernesto Troiano , GFT
- DATAMITE Jordi Arjona, ITI
- PISTIS Yury Glikman, FHG FOKUS

Open discussion

Next steps'



EU DATA+



The EUDATA+ Cluster unites seven projects on data monetisation and exchange, fostered by the European Commission's push for collaboration under programmes like Horizon Europe.

Beyond coordination, it now drives shared goals and knowledge transfer, maximising impact and enhancing the development of data-sharing technologies.

An infographic titled "Meet the projects" on a dark blue background. It features a central white and orange EUDATA+ logo. Surrounding it are seven project logos, each in a rounded square with an orange border: enR|chMyData (a white lotus flower), IPCAST PROJECT (a white plug icon), datamite (a white 'd' with a vertical bar), FAME (the word "FAME" in white), GRAPH MASSIVIZER (a white star icon), PISTIS (a white gear icon), and ExtremeXP (a white circular network icon). At the bottom left is a QR code, and at the bottom right is the text "Scan the QR code to learn more about the EUDATA+ cluster and the projects involved!" and the European Union logo with the text "Funded by the European Union".

DATA LIFECYCLE

DATA PROVIDER

Data Ingestion



Data Enrichment



Data Product Creation and Preparation



Data Publication



DATA EXCHANGE

Data Exchange Environment

DATA CONSUMER

Data Acquisition

Data Consumption



Working groups of the EUDATA+ cluster



TECH GROUP

Scientific and Technical
Collaboration



ECOSYSTEM GROUP

Communication & Dissemination
and Community Building



BUSINESS GROUP

Exploitation, Business topics
& Strategies

Meet the projects



Scan the QR code to learn more
about the EUDATA+ cluster and
the projects involved!

FAME

www.fame-horizon.eu



This Project has received funding from the European Union's Research and Innovation programme under grant Agreement no 101092639

FAME Background and Motivations

EU Data Strategy

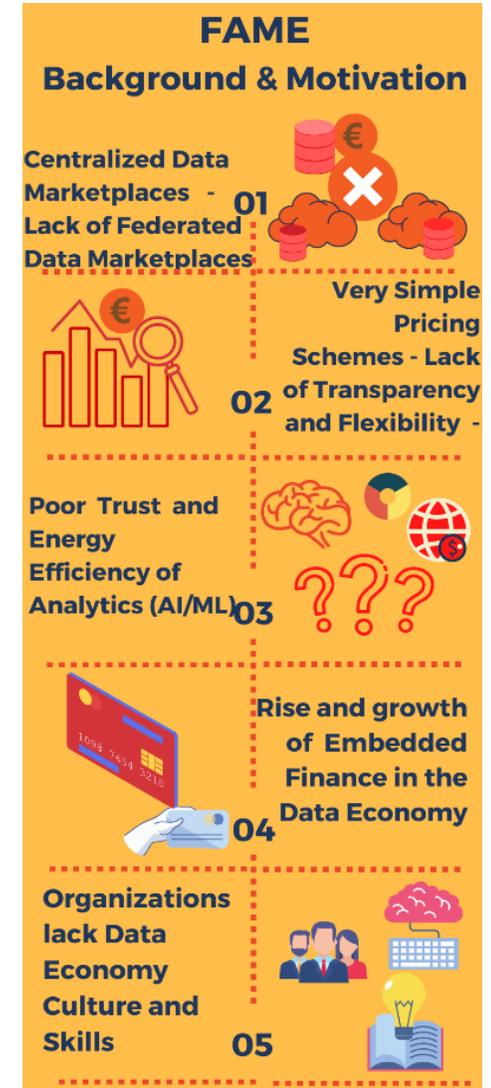
Limitations in Data Assets Discovery, Exchange, Trading

No Flexibility & Lack of Transparency in Monetization of Data Assets

Poor Integrated of Trusted & Energy Efficient Analytics

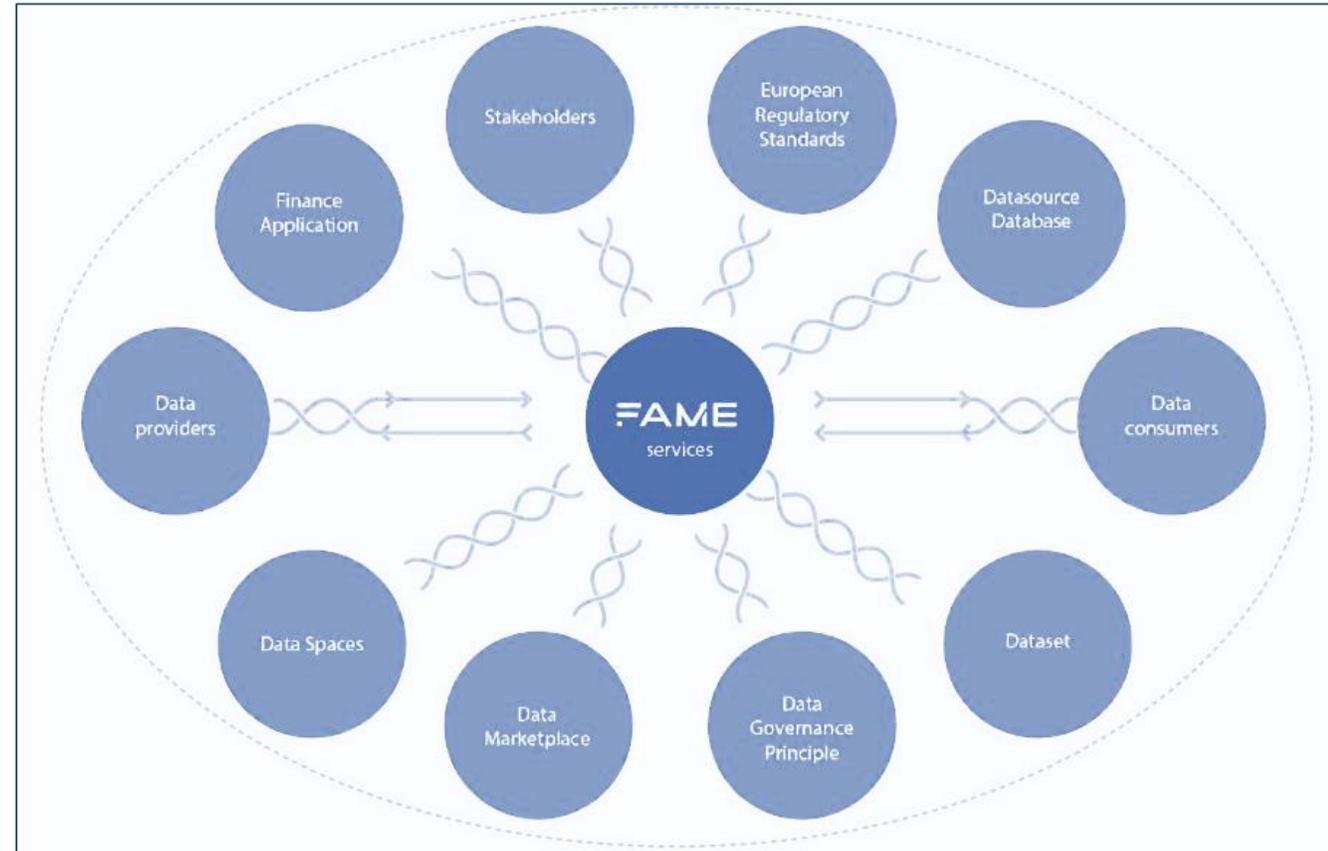
Rise of Embedded Finance & Open Finance Services

EU Complex and articulated regulatory frameworks

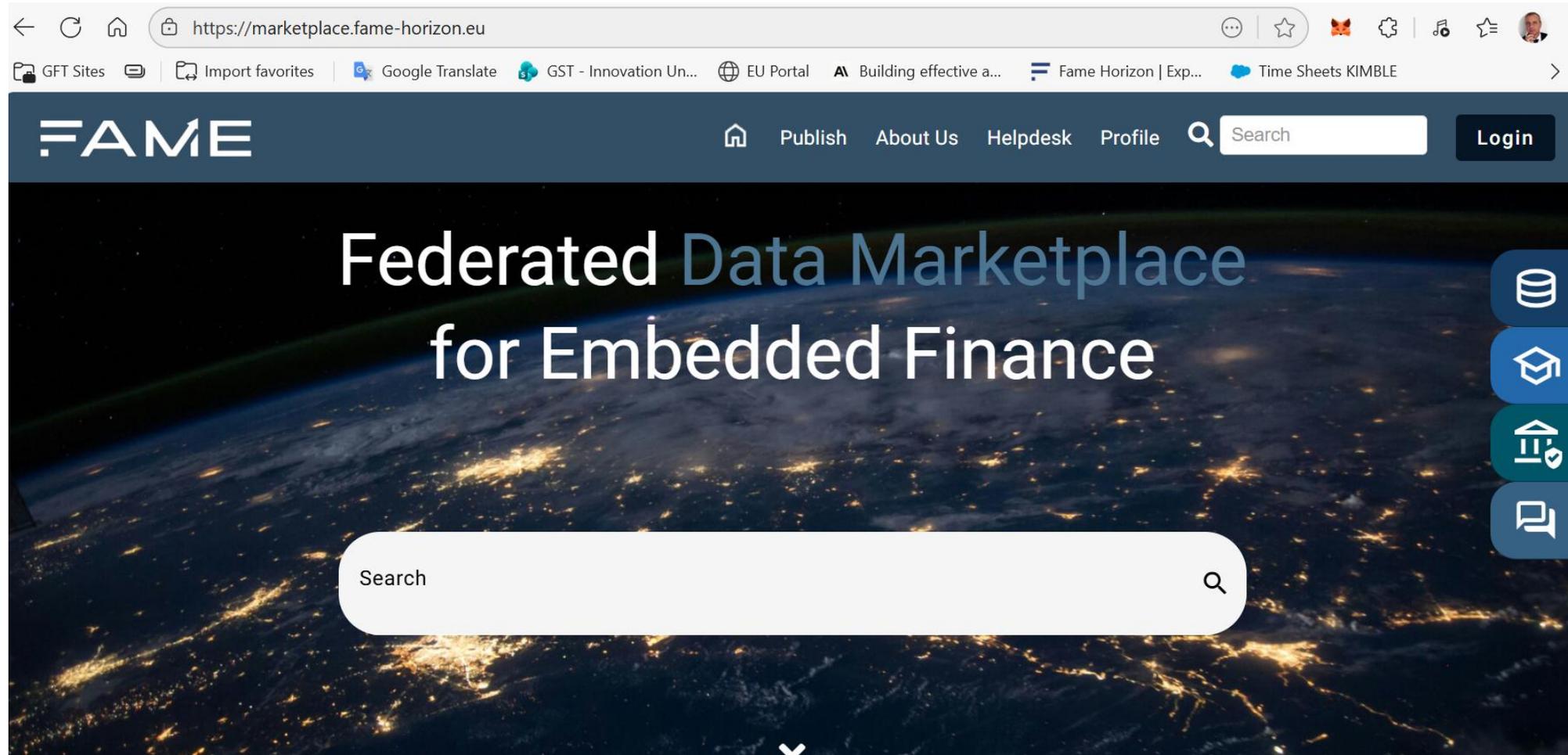


FAME Mission

FAME is on a mission to create a **viable platform and ecosystem** with the ambition to release a reference implementation of a federated **Marketplace** and **Dataspace** for the Financial Sector by embracing the EU data strategy and EU values on data sharing.



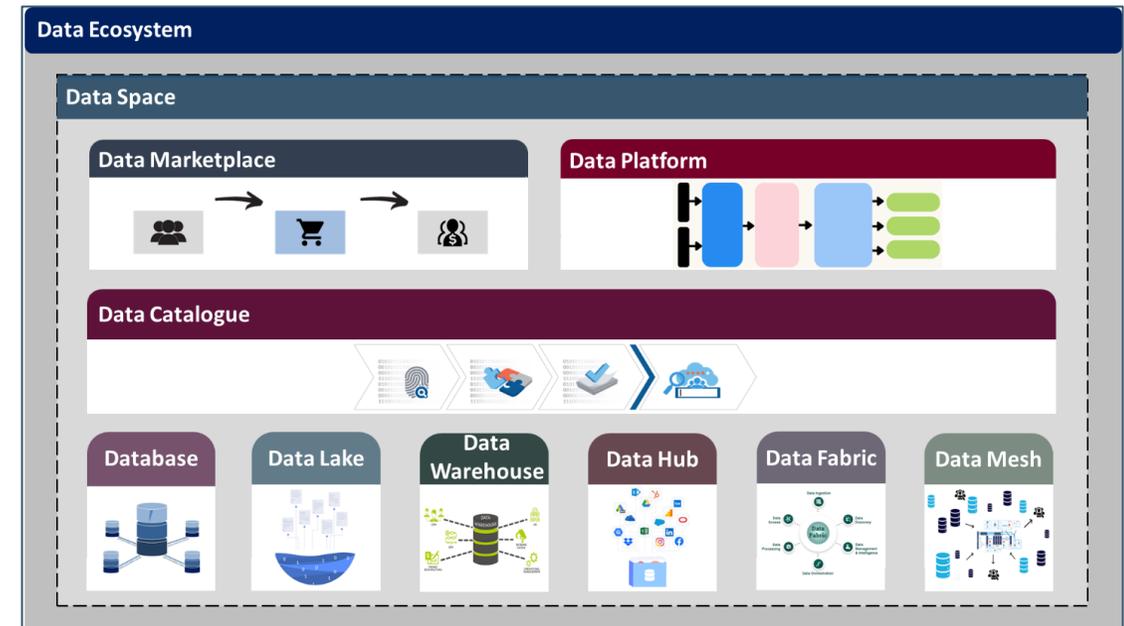
https://marketplace.fame-horizon.eu



FAME, an EU Funded Project

The FAME project has developed the **technology, business** and **legal** frameworks, compliant with emerging **European strategies** and initiatives, to release a unique, open, publicly accessible, trustworthy, energy efficient, and secure **federated data marketplace and platform** for the **finance sector**, featuring:

- Secure, interoperable and regulatory **technologies and solutions**.
- Decentralized, programmable, **trading and pricing** for data assets.
- **Integration of energy efficient** Analytics, ML featuring Explainable AI (XAI), Situation Aware Explainability (SAX).



Traditional Data Marketplaces

✗ Centralized

- └ Data uploaded to platform
- └ Vendor lock-in
- └ Loss of sovereignty

✗ Dataset-Only

- └ Limited to raw data
- └ No AI models or services
- └ No educational content

✗ Opaque Pricing

- └ No market transparency
- └ Guesswork valuation
- └ Platform sets rules

✗ Compliance Burden

- └ Manual verification
- └ Post-publication risk
- └ Multi-jurisdiction complexity

FAME Approach

✓ Federated

- └ Data stays with publisher
- └ No vendor lock-in
- └ Full sovereignty maintained

✓ Multi-Asset

- └ Datasets + AI models
- └ Running services + APIs
- └ Educational content

✓ Transparent

- └ Blockchain-verified history
- └ AI-powered pricing advisory
- └ Democratic governance

✓ Compliance-by-Design

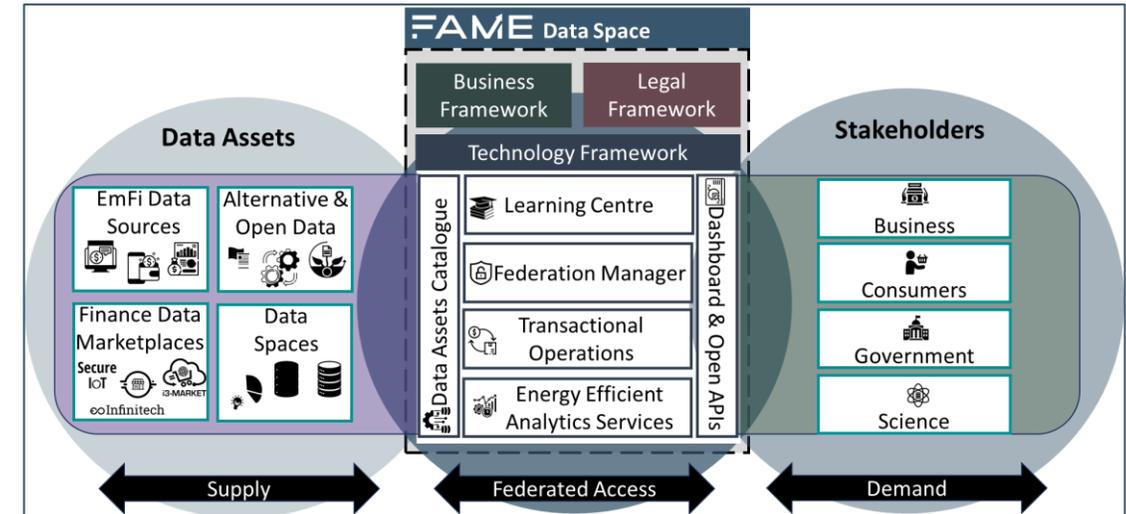
- └ Automated verification
- └ Smart contract enforcement
- └ EU regulations built-in



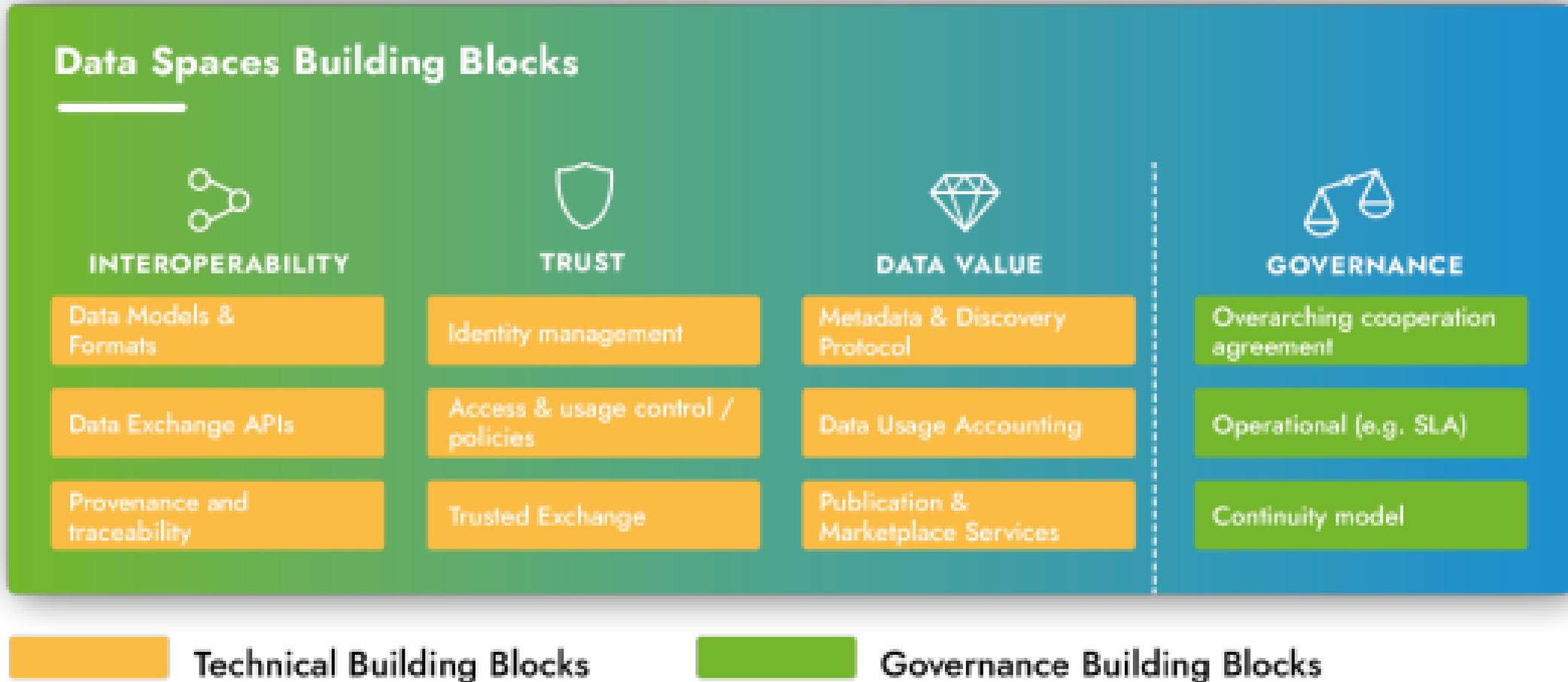
FAME as a Data Space

FAME is a Data Space, adhering to EU regulations for data governance providing:

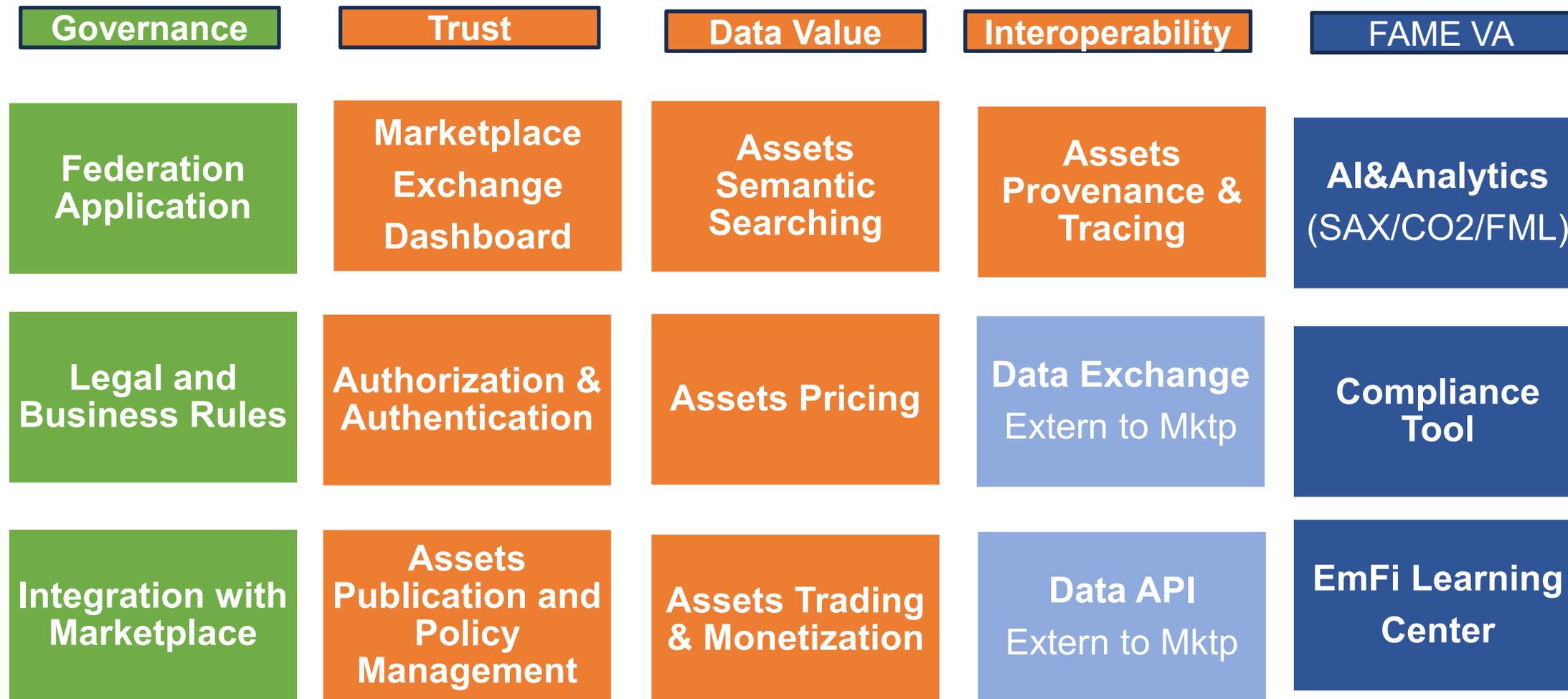
- **Financial Data Marketplace:** FAME provides a dedicated platform for trading data assets specifically suited for the financial sector.
- **Federated Governance:** FAME establishes a federated business and governance model to ensure data sovereignty and trust during exchange within digital ecosystems.
- **Shared Principles:** All data transactions conducted through FAME are grounded on commonly agreed-upon principles



DSSC Blueprint



FAME vs DSSC Blueprint



FAME Unique selling points

Federation, Not Centralization

- Data stays with publisher - No custody transfer – Multi-marketplace interoperability
- Adopts an **extension of DCAT** to include marketplaces and other type of data assets including the above mentioned
- **Example:** Small fintech publishes proprietary scoring model without infrastructure complexity

Business Models Flexibility

- PAYG, PAYU, Subscription - Dynamic pricing
- Automated enforcement - Custom business models
- **Benefit:** Any monetization strategy, automated execution

Blockchain-Based Transparency

- Offering Tokens (ERC-721): Immutable provenance, Transparent pricing history
- Data Access Tokens (ERC-1155)
- Money Tokens (ERC-20)
- **Result:** Verifiable trust, not marketing claims

Reference and Regulatory Frameworks

- Adheres to the **DSSC conceptual model** adding components to enable monetization of assets
- GDPR, Data Act, FIDA , MiCA, PSR, 4AML
- Automated verification – Democratic governance
- **Impact:** Regulatory confidence by design

FAME Business Value

- FAME Marketplace supports a new kind of data economy, federated, transparent, sustainable, and aligned with European values.
- FAME Consortium commitment is to provide to a large community in the finance sector, the resources they need to build applications with trusted, secure, quality data from an ecosystem they contribute to build
- We made it for finance to prove we can make it for everything

FAME is all about publishing data and trust



Our Partners

- Large Industries (Digital Finance/Vendors/System Integrators)
- Financial Institutions, FinTech, EmFi Data Assets Providers
- SMEs, BigData, AI, Blockchain, Vendors, Integrators
- Universities and Research Centers
- FinTech Hubs, Associations & Legal Partners



DSSC Questionnaire

- Results

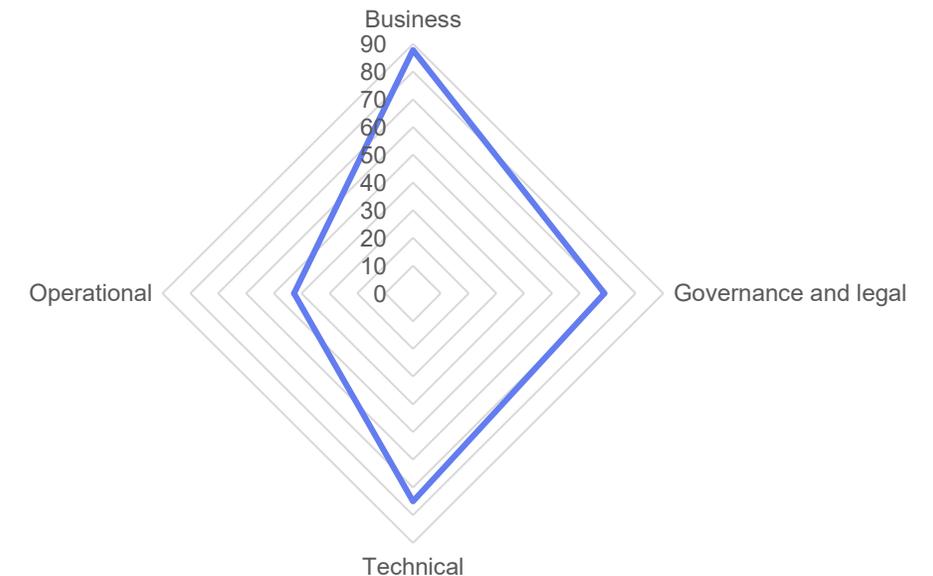


DSSC Maturity Model

FAME SELF-ASSESSMENT

Radar axis	Value (0-100)	Note
Business	87.78	Avg of Q1-Q3
Governance and legal	68.81	Avg of Q4-Q8
Technical	75.00	Avg of Q9-Q17
Operational	42.73	Avg of Q18-Q19

FAME MATURITY MODEL SELF-ASSESSMENT



Suggestions on DSSC Model & Questionnaire

- The DSSC questionnaire fits the purpose
 - It fits as it reflects the four dimensions (business, legal, technical and operation)
 - If more dimensions are added, than the questionnaire should reflect it.
 - Scoring is fine as it is a metric. Any metric is arbitrary, and the one chosen fit the purpose.
- More comments
 - Questionnaire must address different stakeholders, therefore maybe should be highlighted if the questions should be answered by business, legal or technical.
 - Questionnaire is quite long, that's why indicate who should reply to what is important do divide the tasks between different stakeholders
 - Some aspects more relevant to our implementation are missing, but as a generic model is just fine. However, a section with free self-addressing q&a could be added as ancillary notes.



Federated Decentralized Trusted Data Marketplace for Embedded Finance

FAME

www.fame-horizon.eu



This Project has received funding from the European Union's Research and Innovation programme under grant Agreement no 101092639

Agenda

Insights from CEN/CENELEC JTC 25

Learnings from projects 30'

- FAME Ernesto Troiano , GFT
- [DATAMITE](#) Jordi Arjona, ITI
- PISTIS Yury Glikman, FHG FOKUS

Open discussion

Next steps'



datamite

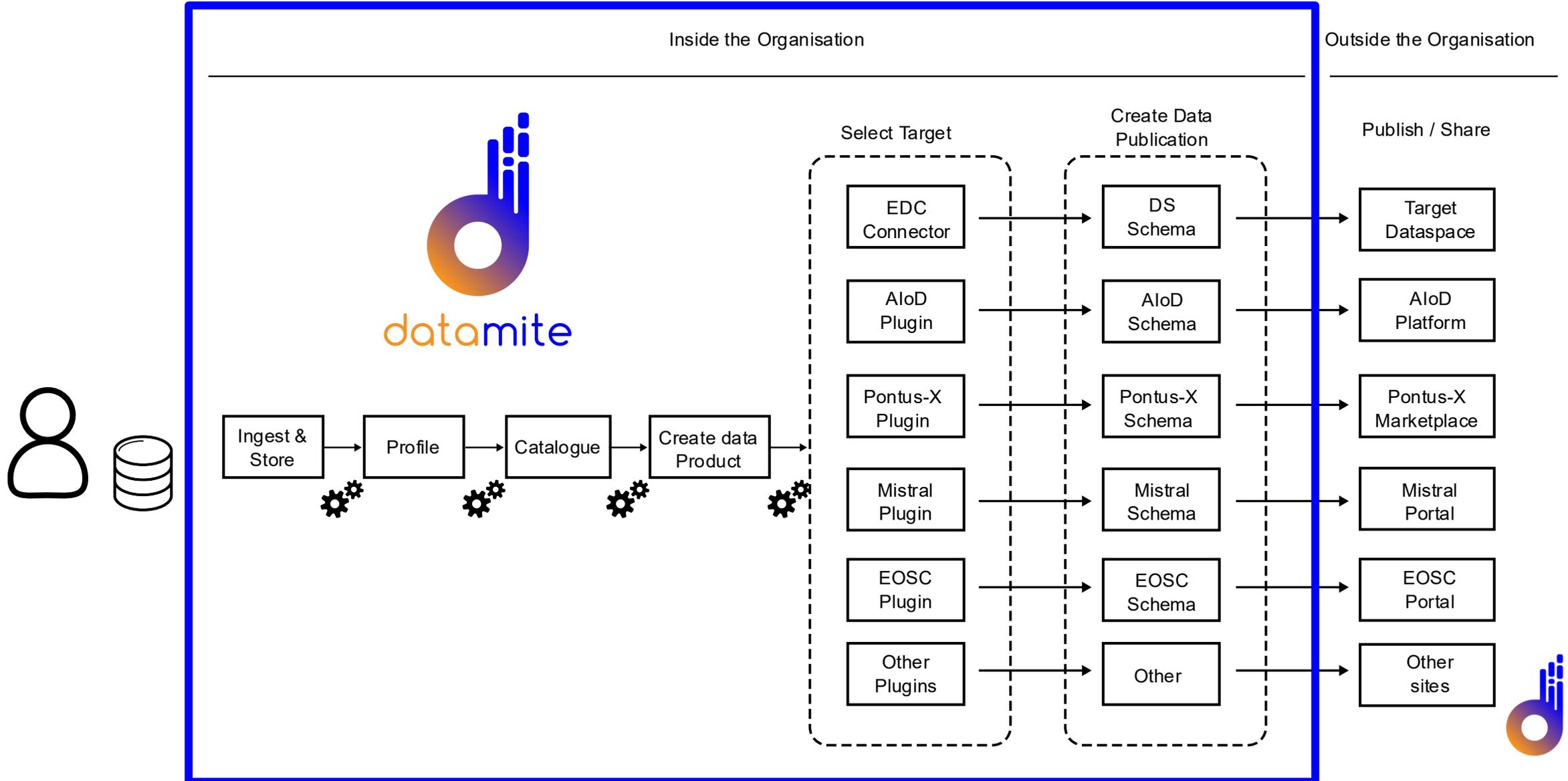
Interoperability, Data governance, Data spaces maturity assessment and the path to Standardisation

EUDATA+ Joint workshop

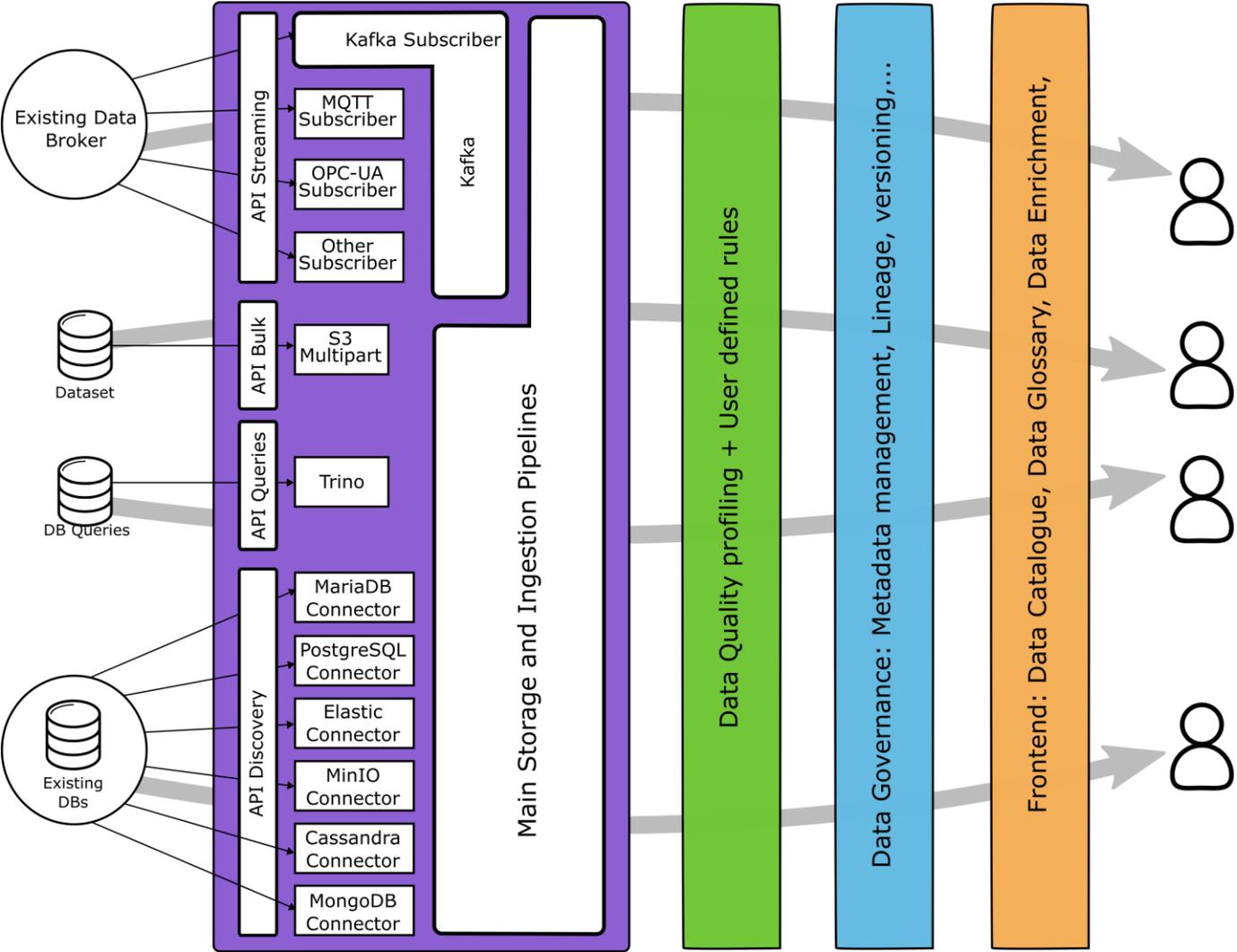


Funded by
the European Union

What is DATAMITE's scope?



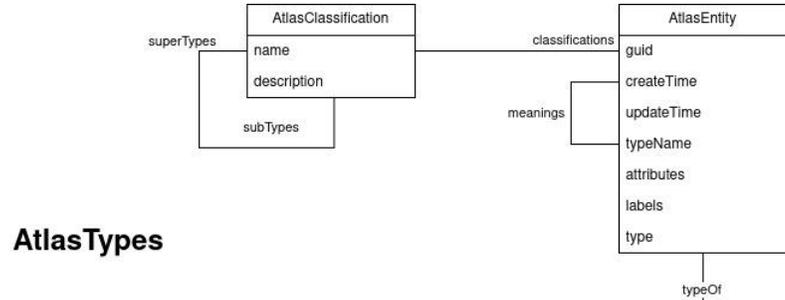
The data onboarding process



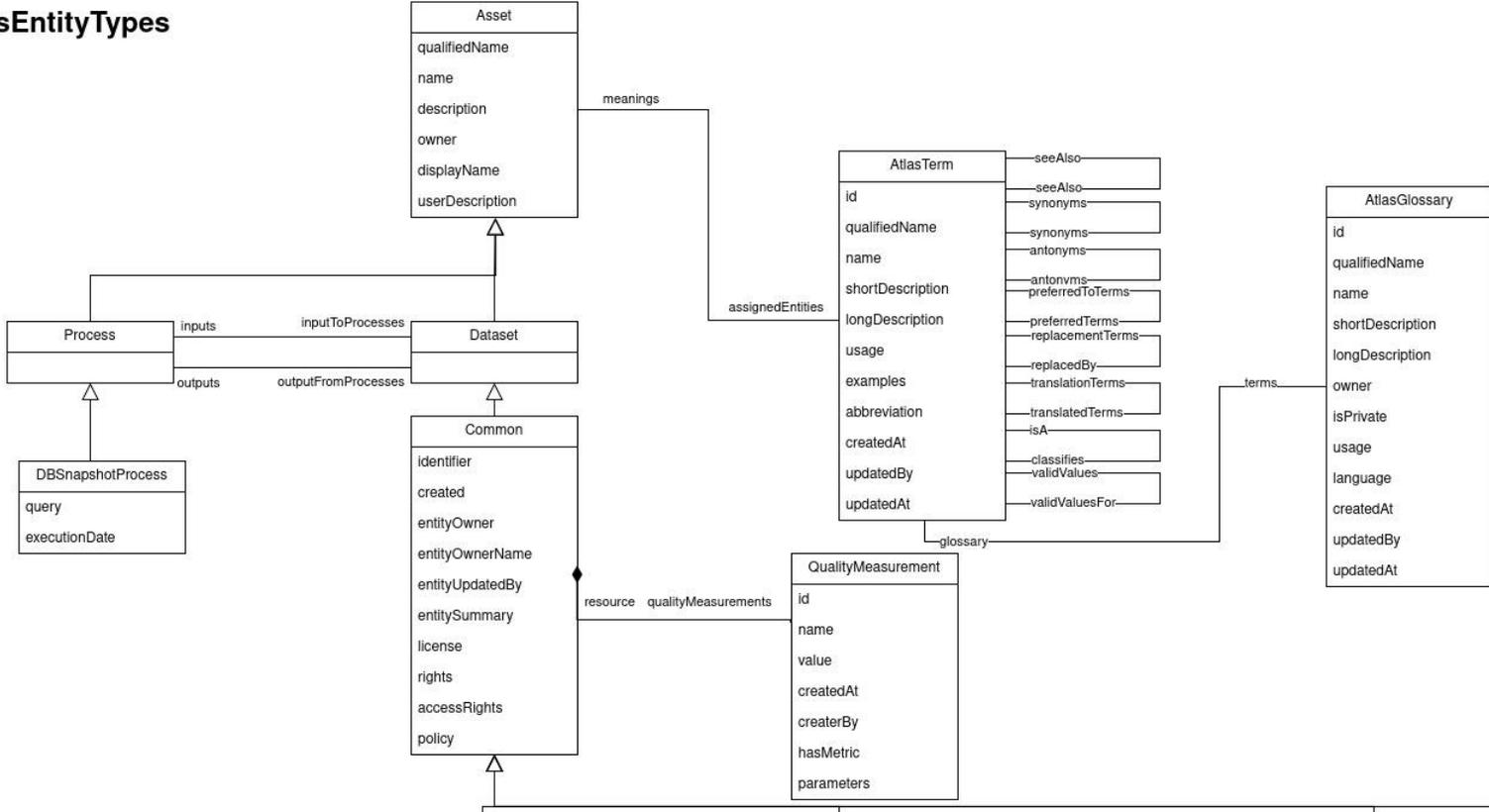
COMPLETELY
INSIDE
THE
COMPANY



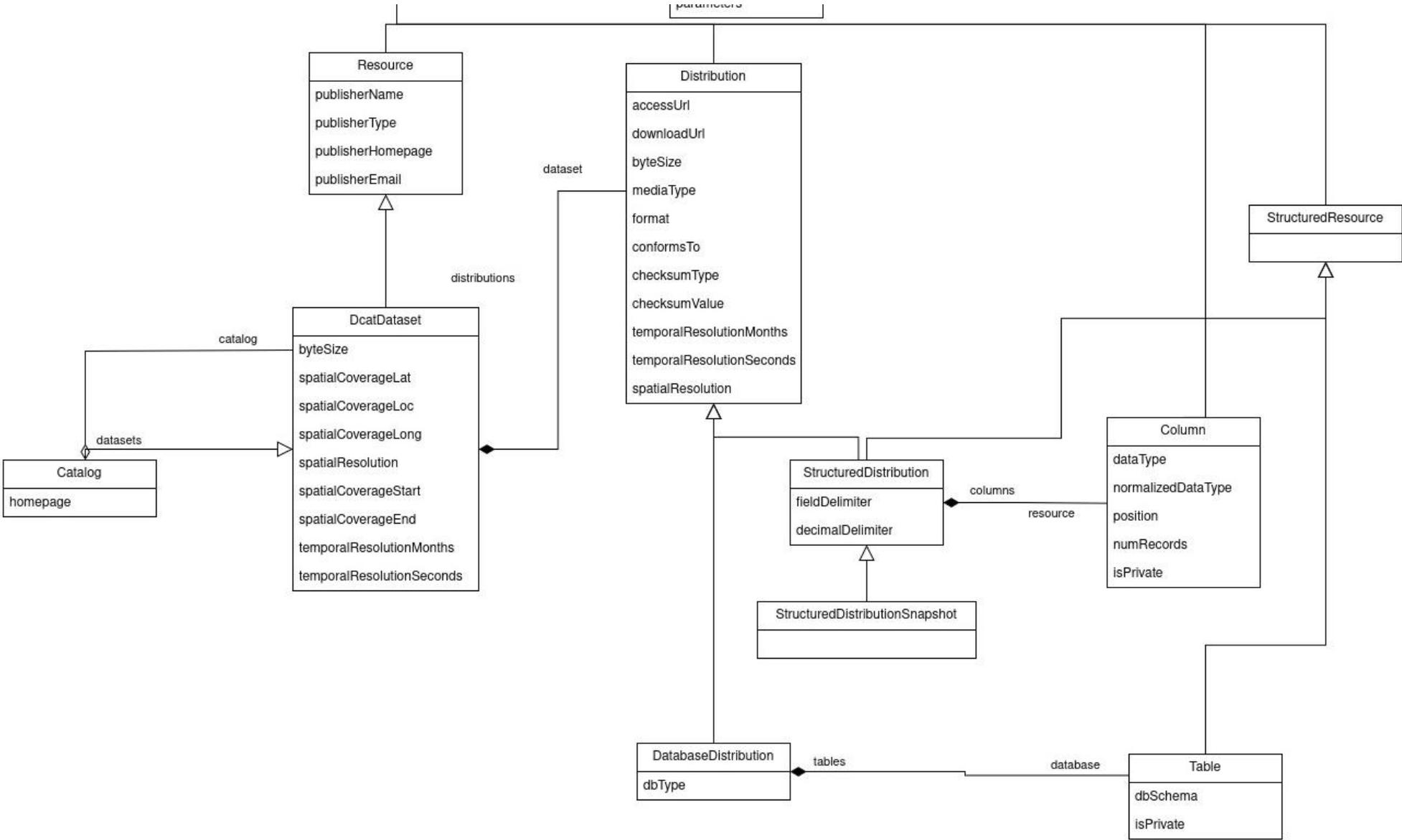
DATAMITE's Metadata Model



AtlasEntityTypes



DATAMITE's Metadata Model



DATAMITE's Metadata Model

- We consider three types of metadata:
 - The one that already exists (technical: column names, owners, creation dates, ...)
 - The one the user provides (business: description, business terms from glossaries, ...)
 - The one we compute (business+technical: quality KPIs, both informative and user-defined)
- Based on DCAT (we also use part of the metadata structure from Atlas)
- DCAT purpose: This is dataset and its distributions, take them.
- Limitations
 - DCAT is insufficient to represent assets within a company. Lacks depth.
 - We have to overload it to include details on columns, quality, etc.
 - Representing Databases is complicated again. We represented them as distributions.
 - Multimedia: what is a distribution? Each file or a zip with all of them?



D-DQV

- Key field: Quality measurement
 - Informs what was measured and what was the result
 - Works for well-known metrics
 - Additional information for dimensions and categories
- Problems
 - There is not a clear standard for dimensions and categories
 - We do not think it has the capability of representing custom metrics

ddqv: Datamite DQV

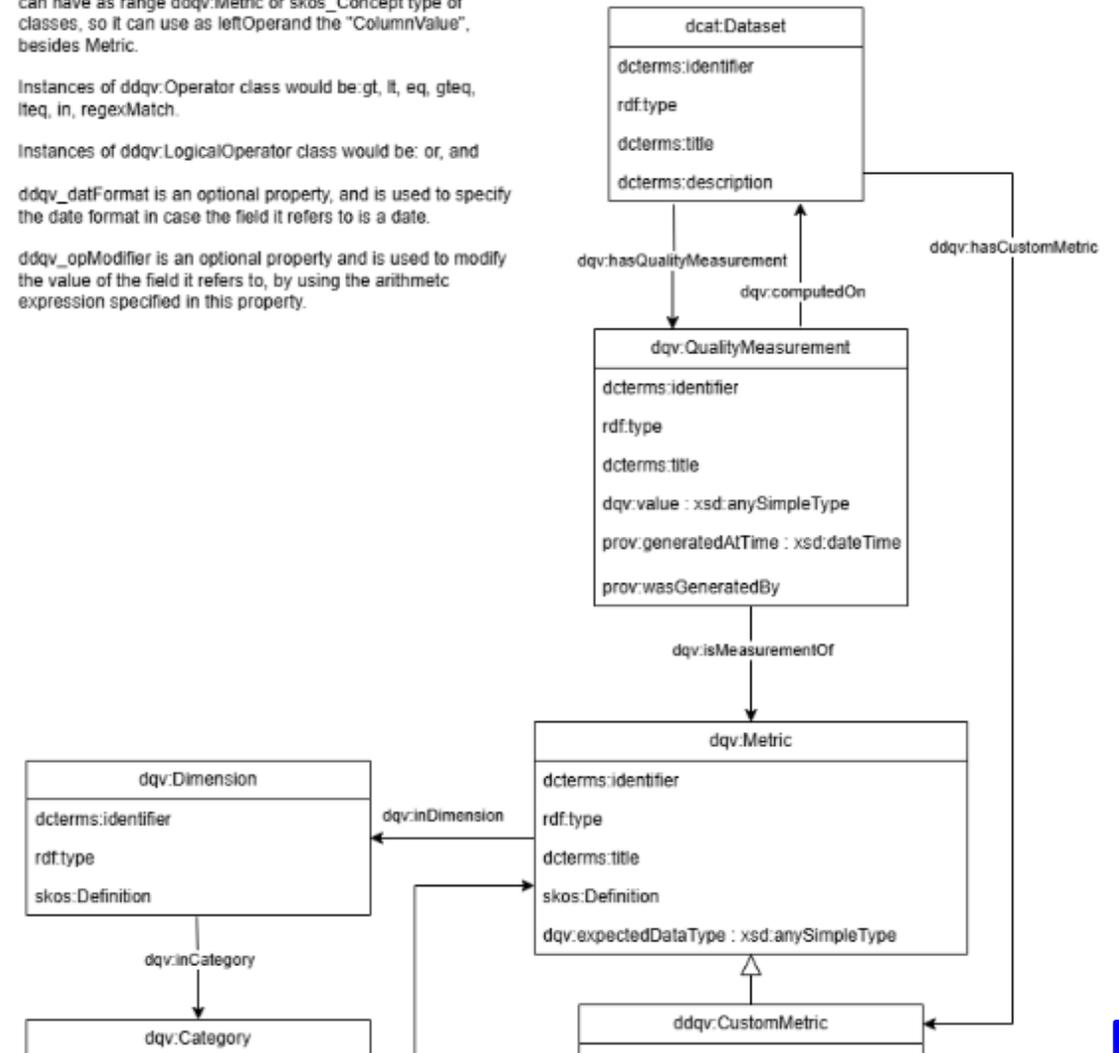
The relationship `ddqv:refersTo` from `ddqv:LeftOperand` class can have as range `ddqv:Metric` or `skos_Concept` type of classes, so it can use as `leftOperand` the "ColumnValue", besides `Metric`.

Instances of `ddqv:Operator` class would be: `gt`, `lt`, `eq`, `gteq`, `lteq`, `in`, `regexMatch`.

Instances of `ddqv:LogicalOperator` class would be: `or`, and

`ddqv_datFormat` is an optional property, and is used to specify the date format in case the field it refers to is a date.

`ddqv_opModifier` is an optional property and is used to modify the value of the field it refers to, by using the arithmetic expression specified in this property.

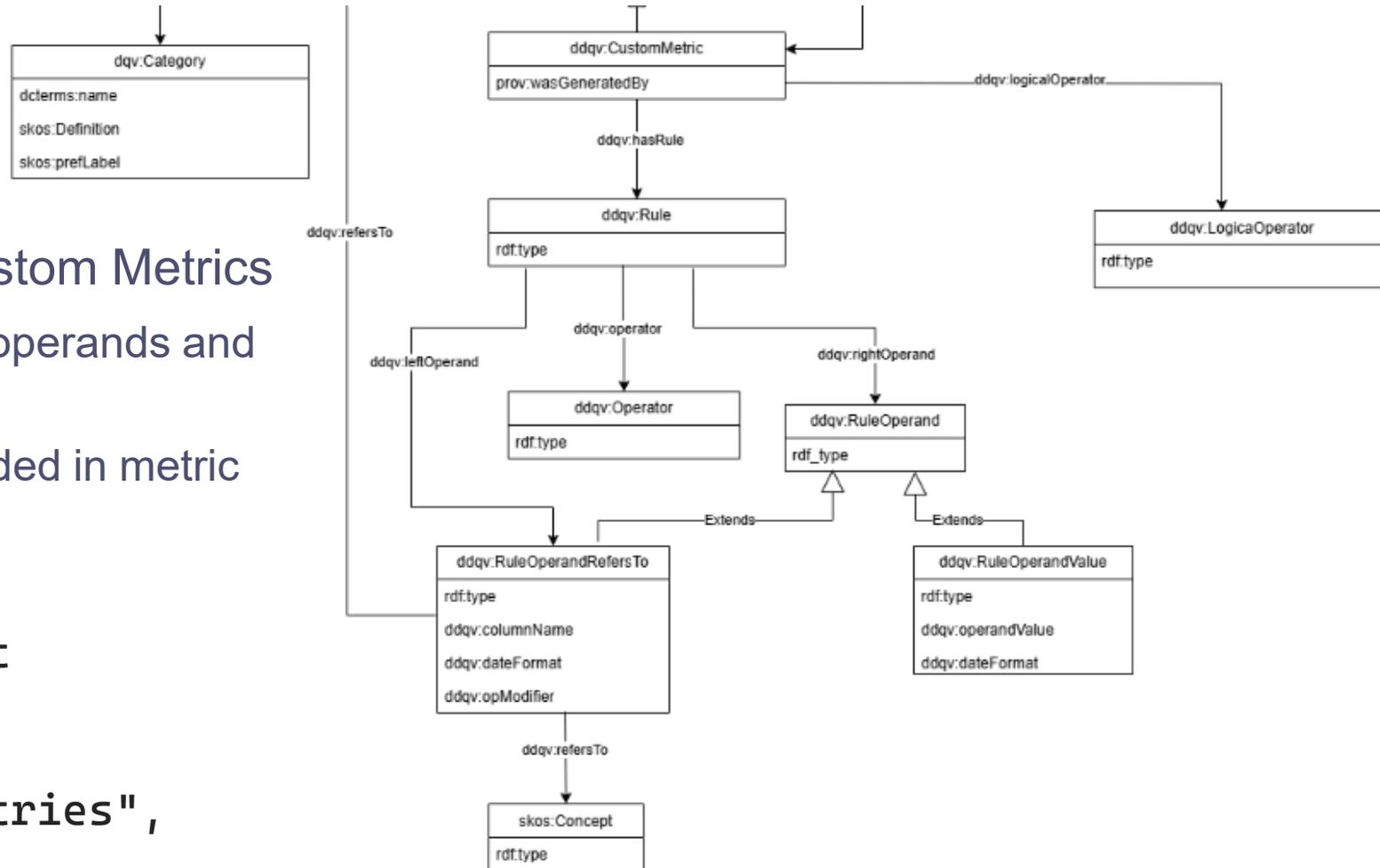


D-DQV

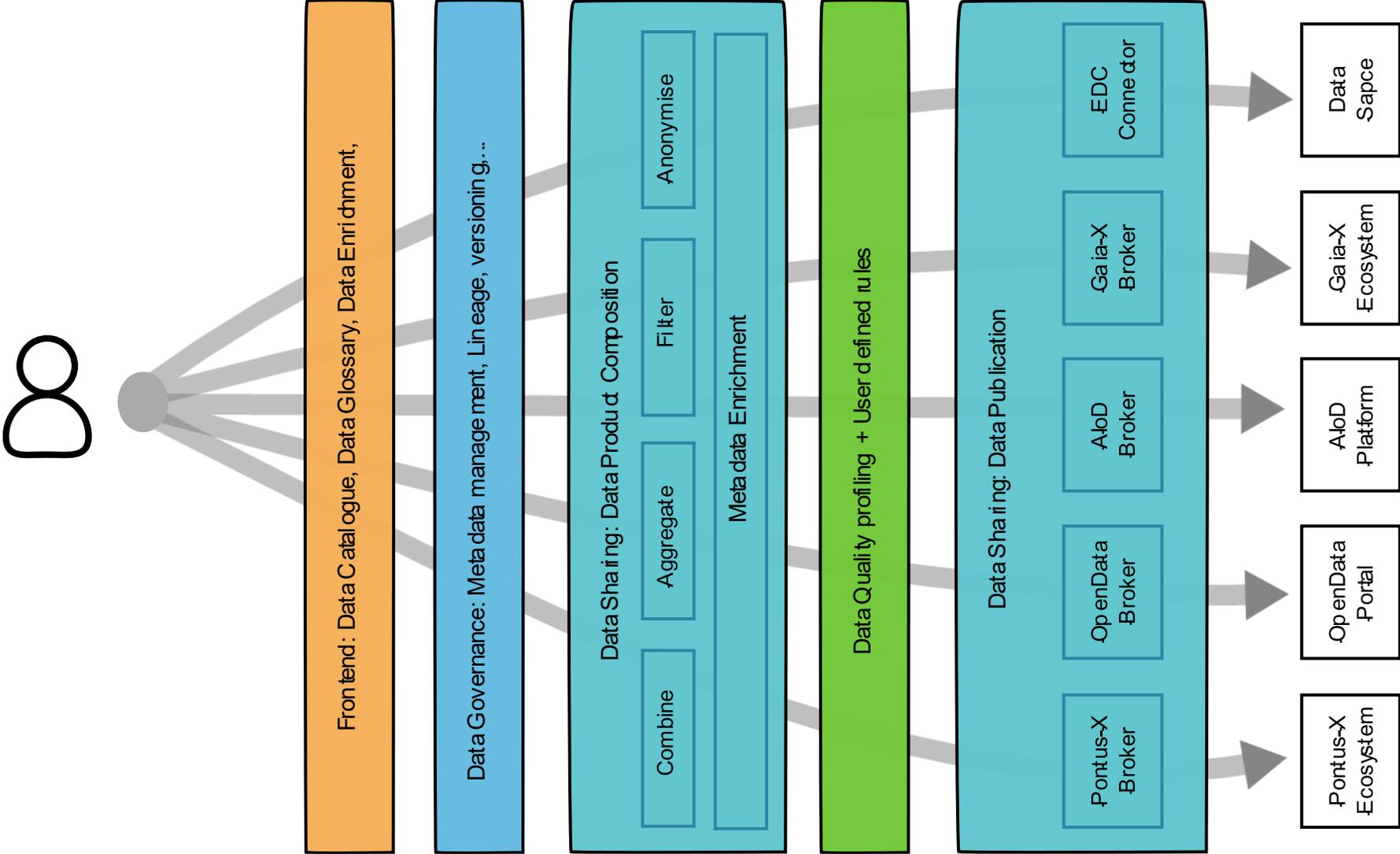
- We extended it to include Custom Metrics
 - Includes information about operands and operators
 - Description is already provided in metric

Example Quality Measurement

```
{  
  "name":  
  "LON_general.duplicated_entries",  
  "metric":  
  "general.duplicated_entries",  
  "value": "8",  
  "createdAt": "2025-11-  
12T09:36:57.257Z",  
  "createdBy": "5bd93f4b-13d9-4567-a44d-  
9baa0ea0062a"  
}
```



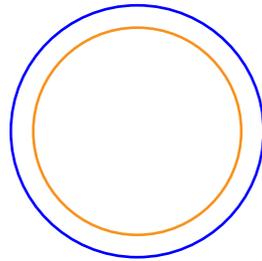
The ingestion process



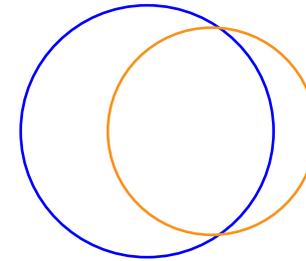
Onboarding of data products

- Plugin-based approach to map from the DATAMITE's Metadata Model to the target ecosystem metadata model.
- Possibilities

- ● DATAMITE
- ● Target Ecosystem



Perfect mapping



Incomplete mapping

- Additionally
 - In the case of Data Spaces -> add data sovereignty policies
 - Manage authentication



Role of Data Quality

- Option a)
 - Evaluate metadata completeness.
 - Data products not fulfilling mandatory fields or below a certain percentage are not informed through the data space catalogue => **That is extortion!** **Yes. It is.**
- Option b)
 - Request information about data quality.
 - How? DQV may be the way.
 - What? Depends on the business case.
 - Requesting Informative KPIs is easy, may not be that valuable
 - Requesting Business KPIs is complex, may require providing extensions for the connector to compute them
 - How to enforce it? I don't think you can. **Extortion again.**





datamite

Thanks!

Keep in touch!



@DATAMITE_EU



/datamite



@datamite_eu



datamite_eu



Funded by
the European Union

Agenda

Insights from CEN/CENELEC JTC 25

Learnings from projects 30'

- FAME Ernesto Troiano , GFT
- DATAMITE Jordi Arjona, ITI
- PISTIS Yury Glikman, FHG FOKUS

Open discussion

Next steps'

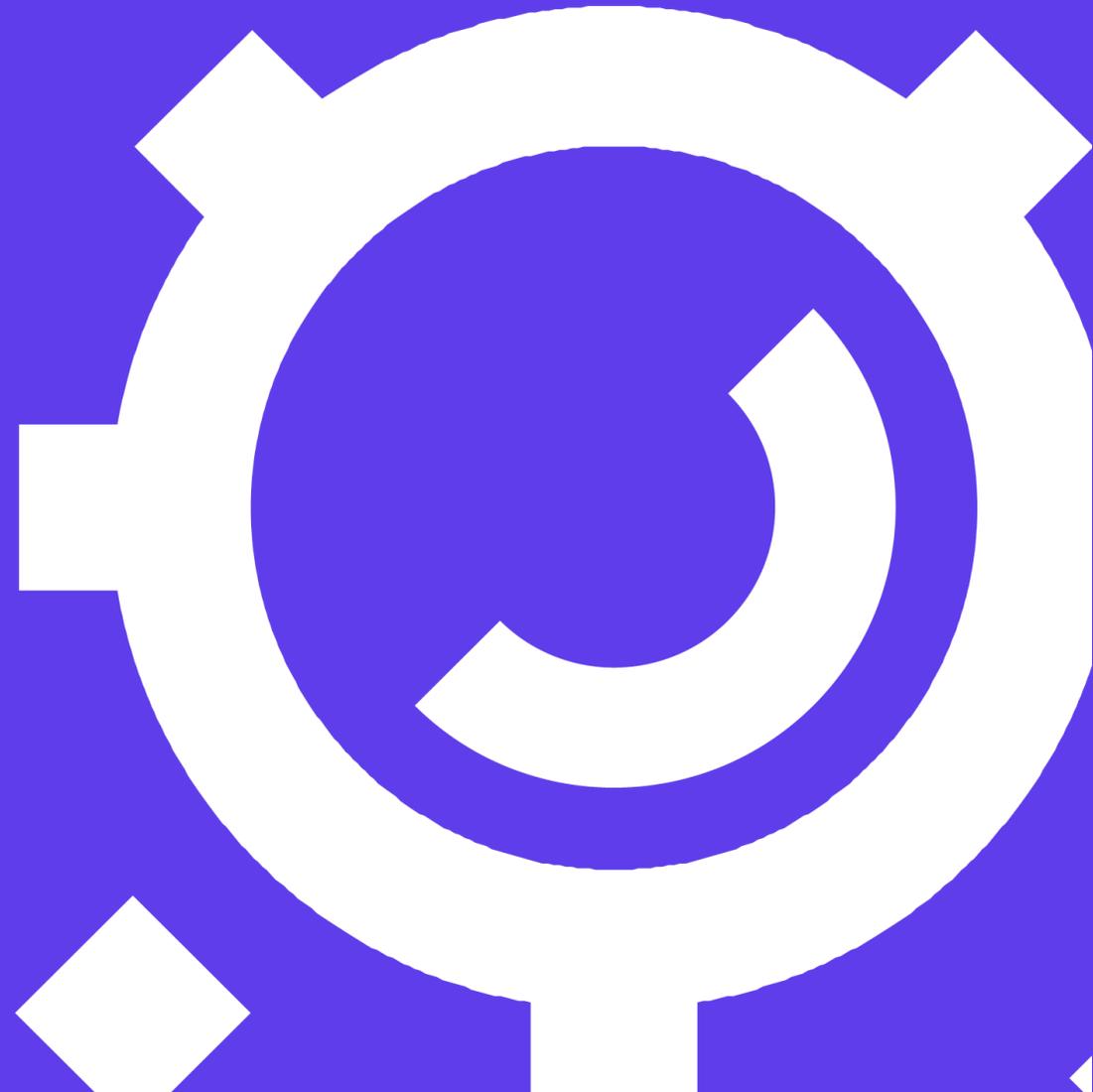
PISTIS and Standards

Dr Yury Glikman
PISTIS Project Coordinator (Fraunhofer FOKUS)

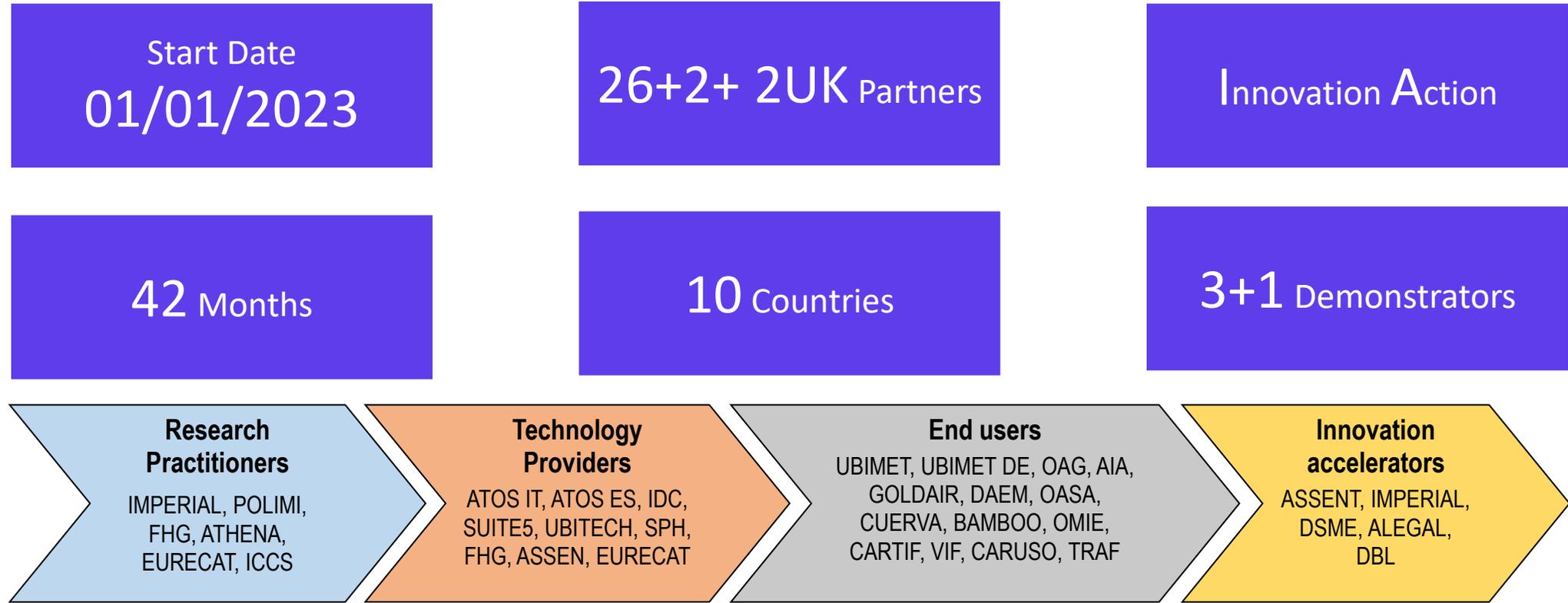
November 2025



Funded by the
European Union



PISTIS Project Basic Facts



PISTIS Aims

- Enable Organisations to **generate more value** from the data.
- Provide the **tools to allow organisations break their silos**, in a trusted, mutual benefiting manner, opening up new business opportunities



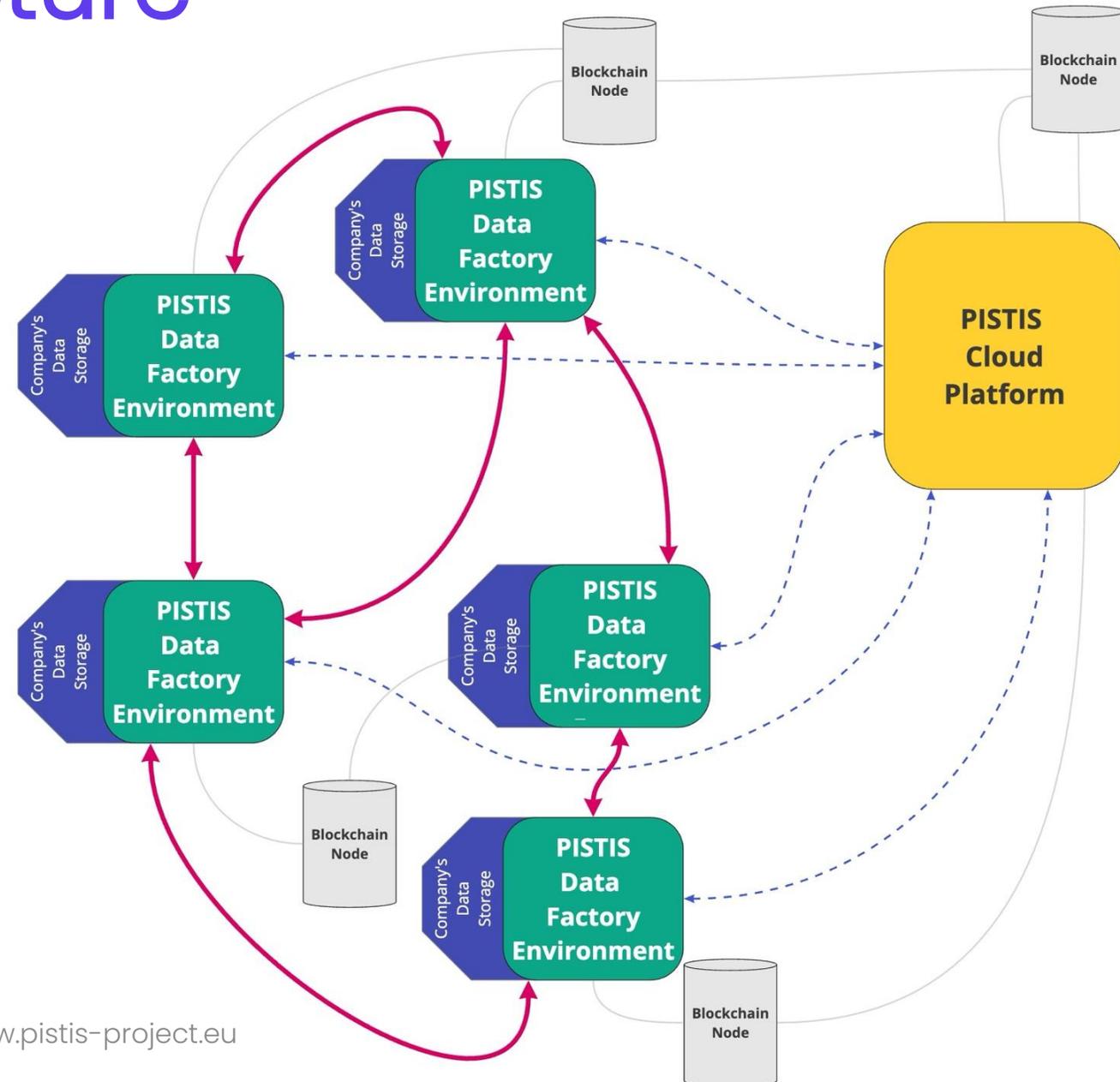
A federated data trading and monetisation platform for secure, trusted and controlled exchange and usage of data assets and data-driven intelligence (derivative data assets)

Reference Architecture

Macroscopic View

2 Core Parts

- PISTIS Data Factory
 - Deployed “locally” over each organisation
 - Holds all the Data of an organisation
- PISTIS Cloud Platform
 - Coordinates the Interaction between the PISTIS Data Factories
 - Holds the central “marketplace” catalogue
 - Oversees all data trading transactions and contracts



What is PISTIS and what it is not

PISTIS is

- ❑ Dataspace and Data Monetisation Platform for Organisations, Out-of-the-Box

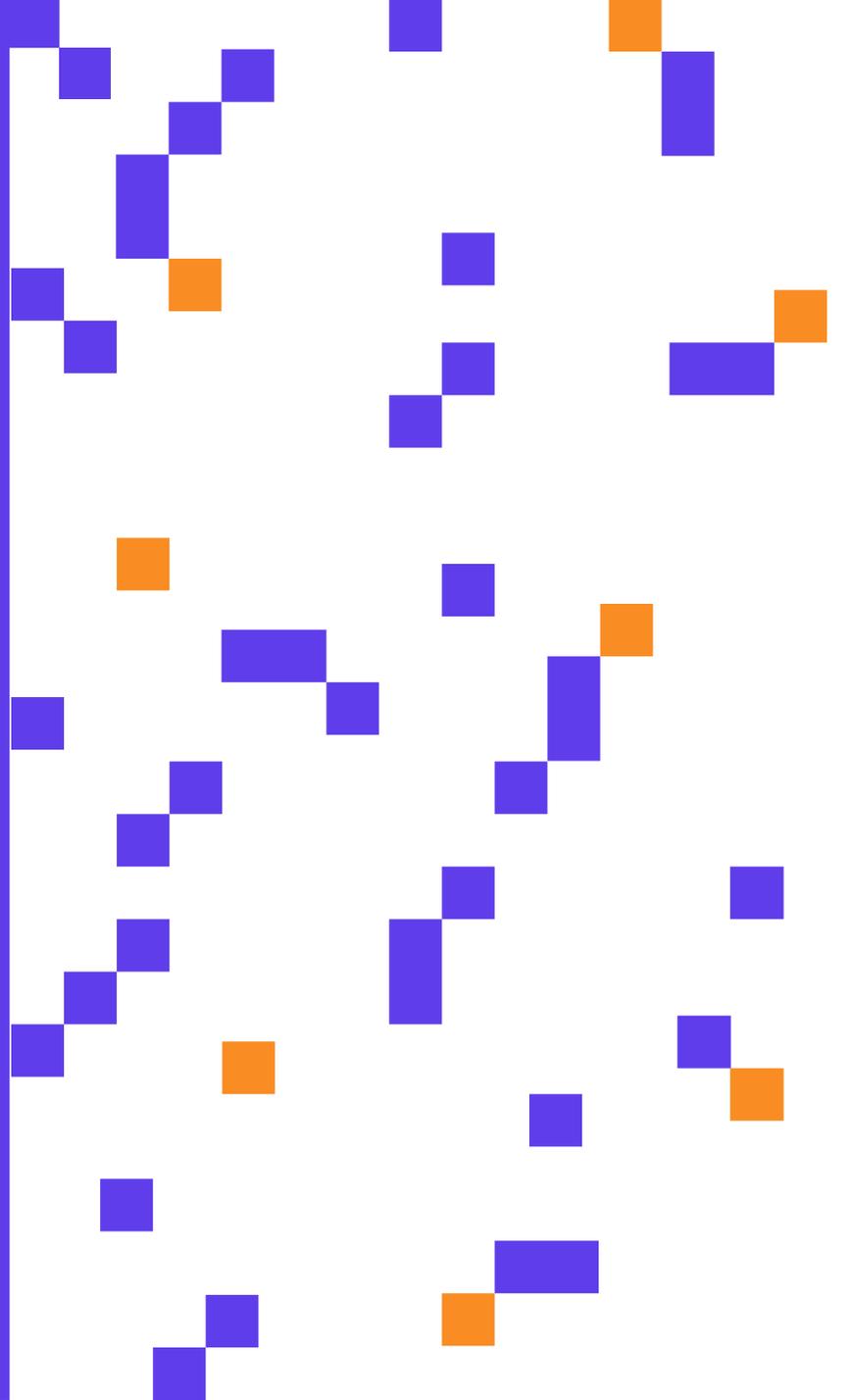
And offers...

- ❑ Data Value Assessment Methods
- ❑ Transactions based on Blockchain technology
- ❑ Contract Monitoring and Management services
- ❑ Interconnection with in-house Data "lakes"
- ❑ A "federated" Catalogue of Data
- ❑ A Data Treatment and Transformation facility

PISTIS is NOT

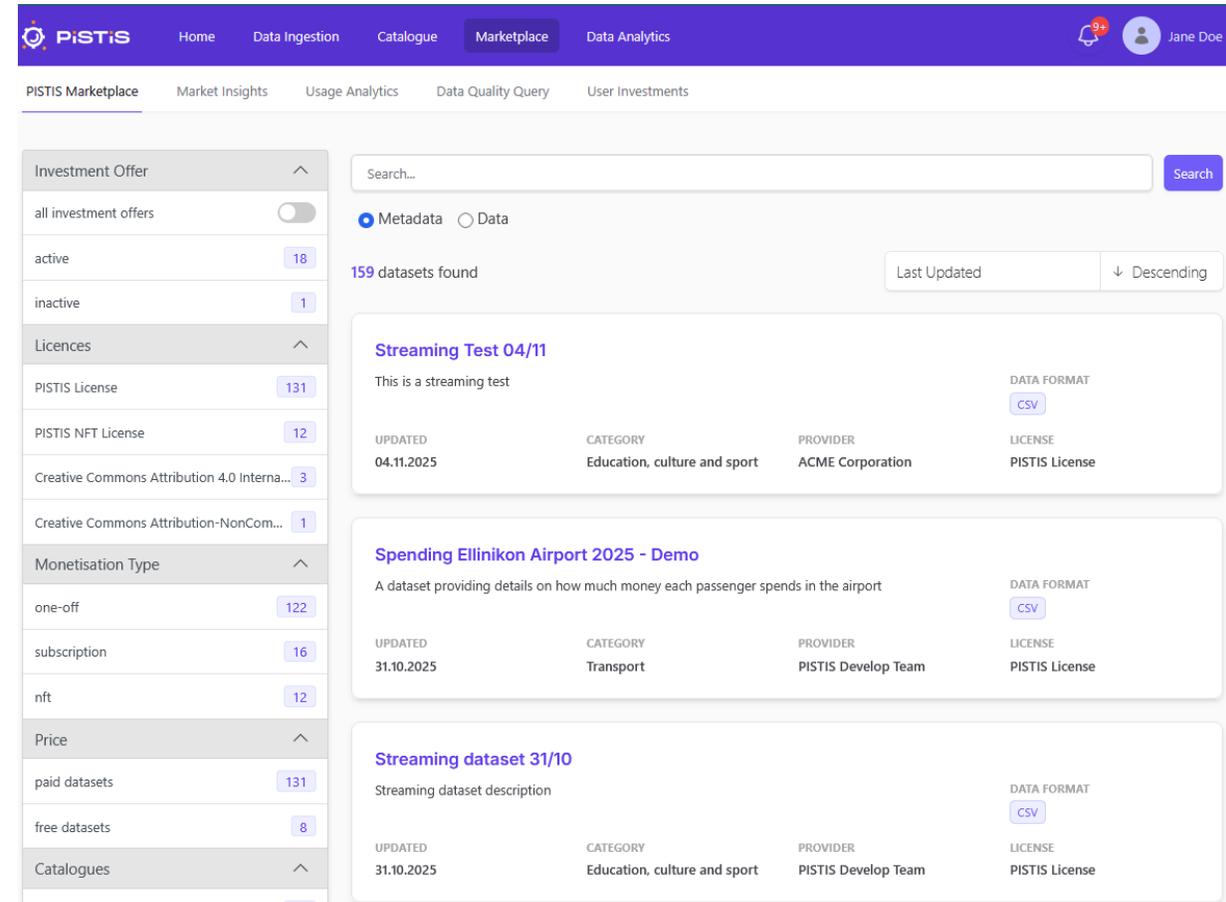
- ❑ A centralised Data Storage
- ❑ A Big Data Platform
- ❑ An Analytics Platform
- ❑ A real crypto market/NFT exchange desk
- ❑ A platform to "host" other applications (e.g. no deep integration with third party apps)

Metadata Management



Metadata Management in PISTIS

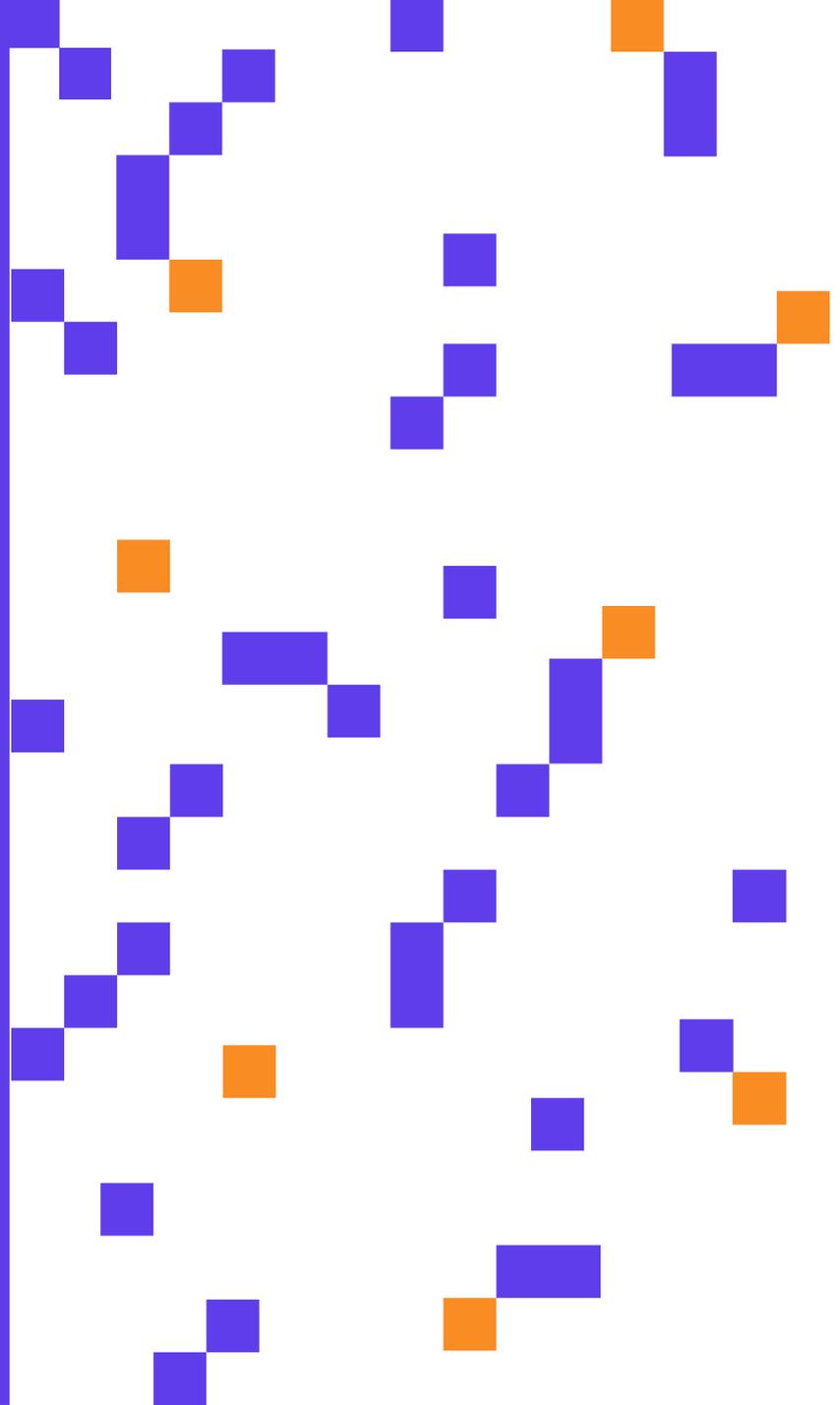
- ❑ Data Model: Custom DCAT-AP profile
- ❑ Implementation based on 
 - ❑ DCAT-AP extension using SHACL
 - ❑ Metadata is managed as RDF in a Triplestore (Virtuoso)
 - ❑ Apache 2.0 License



The screenshot displays the PISTIS Marketplace interface. The top navigation bar includes 'Home', 'Data Ingestion', 'Catalogue', 'Marketplace', and 'Data Analytics'. The main content area shows a search bar, a filter for 'Metadata' (selected) and 'Data', and a search button. Below the search bar, it indicates '159 datasets found' and provides sorting options: 'Last Updated' and 'Descending'. The interface lists several datasets, each with a title, description, and metadata table.

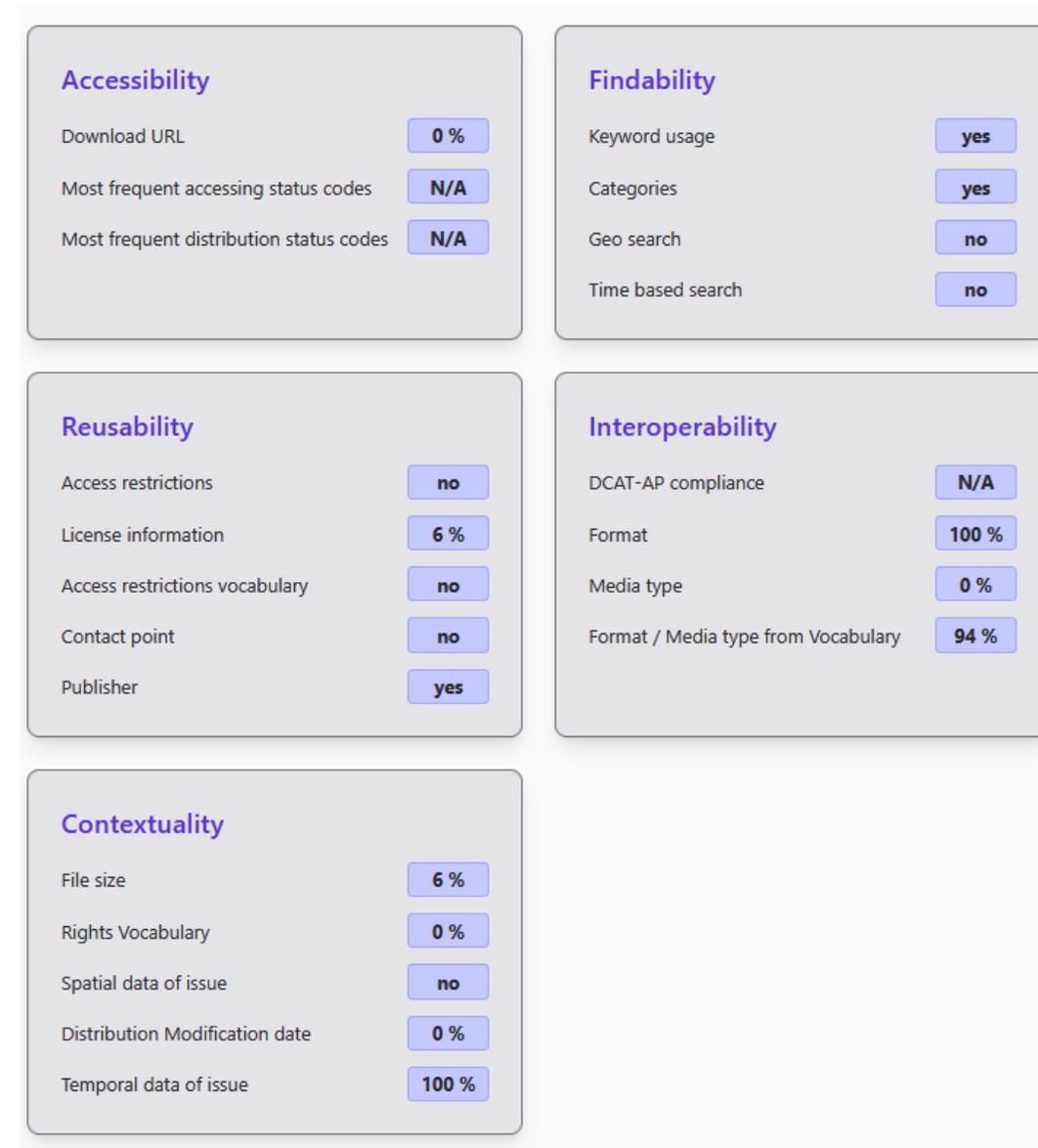
UPDATED	CATEGORY	PROVIDER	LICENSE
04.11.2025	Education, culture and sport	ACME Corporation	PISTIS License
31.10.2025	Transport	PISTIS Develop Team	PISTIS License
31.10.2025	Education, culture and sport	PISTIS Develop Team	PISTIS License

Metadata & Data Quality



MetaData Quality

- ❑ Dimensions are derived from the FAIR principles
- ❑ Measurements are based on the Data Quality Vocabulary (DQV)
- ❑ Measurements on Dataset & Distribution level
- ❑ Compliance with DCAT-AP and DCAT-AP derivatives (PISTIS DM)
- ❑ Implementation based on *piveau-metrics*
 - ❑ Analysis results are stored as RDF in a triplestore



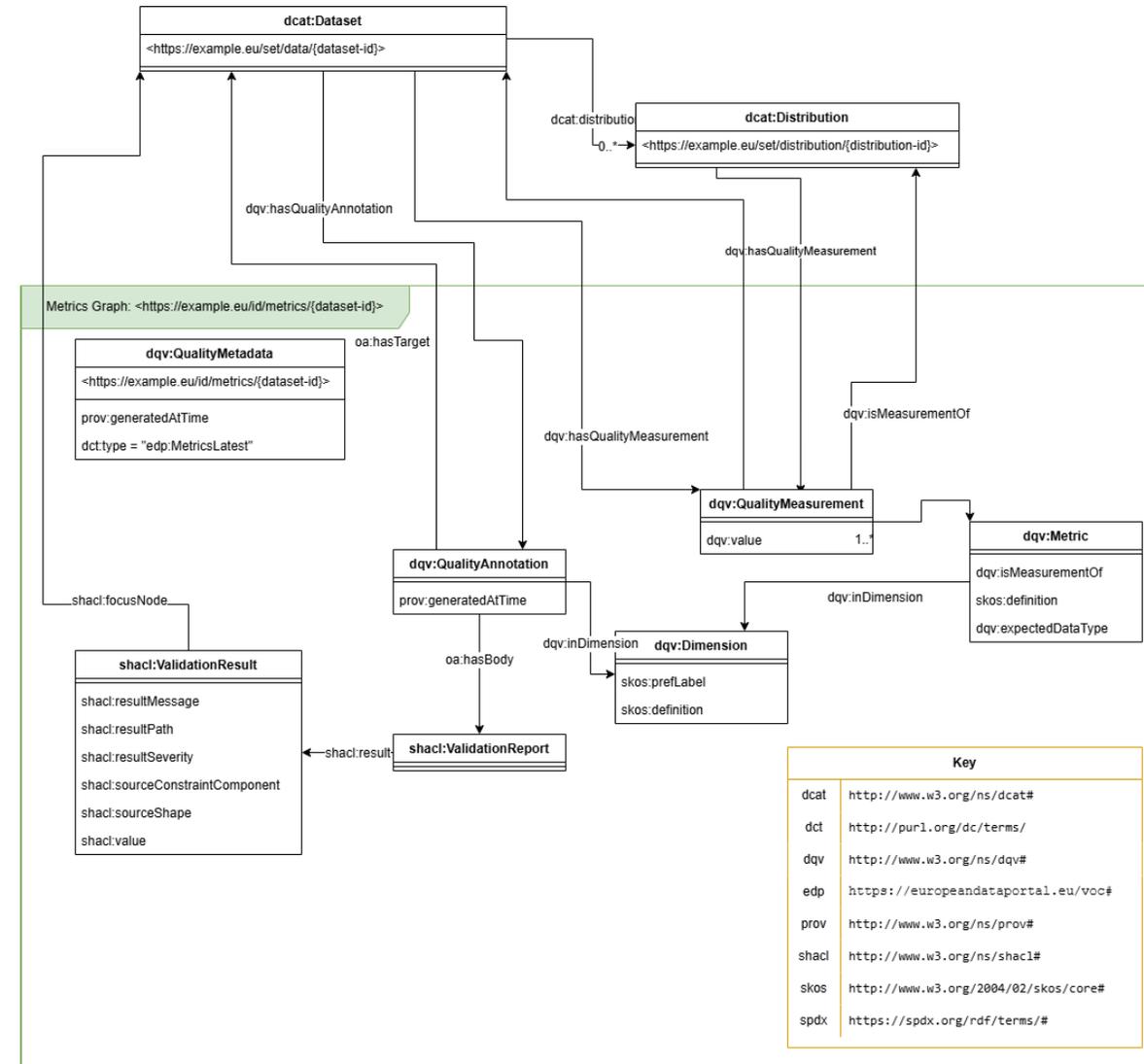
Role of Standards in Metadata Quality

DQV vocabulary

- Each test is a `dqv:QualityMeasurement`
- Each type of test is a `dqv:Metric`
- Metrics are grouped by `dqv:Dimension` (FAIR+C)

SHACL validation

- To validate adherence to the defined schema, metadata is validated via SHACL shapes
- Violations are recorded as `shacl:ValidationResult`
- Linked through to dataset through `dqv:QualityAnnotation`



Data Quality

- Create assessments using **Great Expectations** framework
 - This is a **rule-based engine**
 - Each rule represents an **assertion of how the data should look**
 - There are a set of defined “expectation” types (dqv:Metric)
 - PISTIS maps each “expectation” to the dimension they describe

Accuracy - Defines acceptable values for a feature

Consistency - Defines logical constraints over a feature or between multiple features

Completeness - Defines expectations for where values should exist

Uniqueness - Defines constraints on duplicate values

Format Validity - Defines expectations of distribution structure, i.e. data typing, feature counts, etc.

Test dataset after enrichment process

Accuracy
ExpectColumnValuesToBeBetween (6) ▾

Consistency <- dqv:Dimension
ExpectColumnMinToBeBetween (6) ^ <- dqv:Metric

column name: licence_year	column value: 1.0
column name: licence_protocol	column value: 1.0
column name: licence_type <- dqv:QualityMeasurement	column value: 1.0
column name: district_code	column value: 1.0
column name: address_code	column value: 1.0
column name: area_code	column value: 1.0

ExpectColumnMaxToBeBetween (7) ▾
ExpectColumnStdevToBeBetween (6) ▾
ExpectColumnValueLengthsToEqual (2) ▾
ExpectColumnValueLengthsToBeBetween (5) ▾

Completeness
ExpectColumnValuesToNotNull (14) ▾

Validity
ExpectColumnValuesToBeOfType (14) ▾
ExpectTableColumnsToMatchOrderedList 1.0
ExpectTableColumnsToMatchSet 1.0
ExpectTableColumnCountToEqual 1.0
ExpectTableRowCountToEqual 1.0

Data Enrichment

- ❑ Mapping data types to *owl:DatatypeProperty* from domain ontologies
 - ❑ Generating data representation in a relational data
 - ❑ Data harmonisation
 - ❑ Data versioning
 - ❑ Data quality and vc assessment
- ❑ Description of data structure and semantic
 - ❑ Queryable representation of data



<input checked="" type="checkbox"/> IATA string	<input checked="" type="checkbox"/> ICAO string	<input checked="" type="checkbox"/> Location string
iata city_IATA_Code ⓘ string	DABB	Annaba
iataCarrierCode ⓘ string	EKYT	Aalborg
iataAirportCode ⓘ string	OMAL	AlAin
AircraftType-iataIdentifier ⓘ string	URKA	Anapa
AAR	EKAH	Aarhus
...

Data Lineage

- ❑ W3C PROV-O based
- ❑ Managed as RDF in a Triplestore
- ❑ Records integrity via cryptographic hash of the dataset's lineage in a blockchain
- ❑ Data versions comparison

The screenshot displays the Pistis Data Lineage interface. On the left, a 'Dataset Family Tree' shows two nodes: A.1 (Oct 28) and A.2 (Nov 6), connected by a vertical line. On the right, a table lists lineage activities. The table has columns for VERSION, ID, ACTIVITY, USERNAME, and TIMESTAMP. The first entry (A.1) is 'DATASET REGISTERED' and the second (A.2) is 'DATASET UPDATED'. The 'DATASET UPDATED' activity includes a detailed list of schema changes.

VERSION	ID	ACTIVITY	USERNAME	TIMESTAMP
A.1	33b09883-29ca-43bd-b5e9-8dae9a787d24	DATASET REGISTERED Dataset was successfully ingested via the Job Configurator and registered in the Factory Catalog.	/PDT pdt-01	October 28, 2025 at 03:59:56 PM
A.2	e1e0ccdf-cb1c-4cc6-98f4-369869711c8d	DATASET UPDATED SCHEMA CHANGES: Added 16 column(s): vehicle_id, marker_position, start_position, start_after, start_altitude, arrival_actual_ingate_datetime_utc, duration, arrival_actual_ingate_datetime_local, event_details, event_type, driver_id, event_id, end_altitude, end_after, trip_id, end_position + vehicle_id + marker_position + start_position + start_after + start_altitude + arrival_actual_ingate_datetime_utc + duration + arrival_actual_ingate_datetime_local + event_details + event_type + driver_id + event_id + end_altitude + end_after + trip_id	/PDT pdt-01	November 6, 2025 at 03:58:54 PM

Standardization in Data Quality

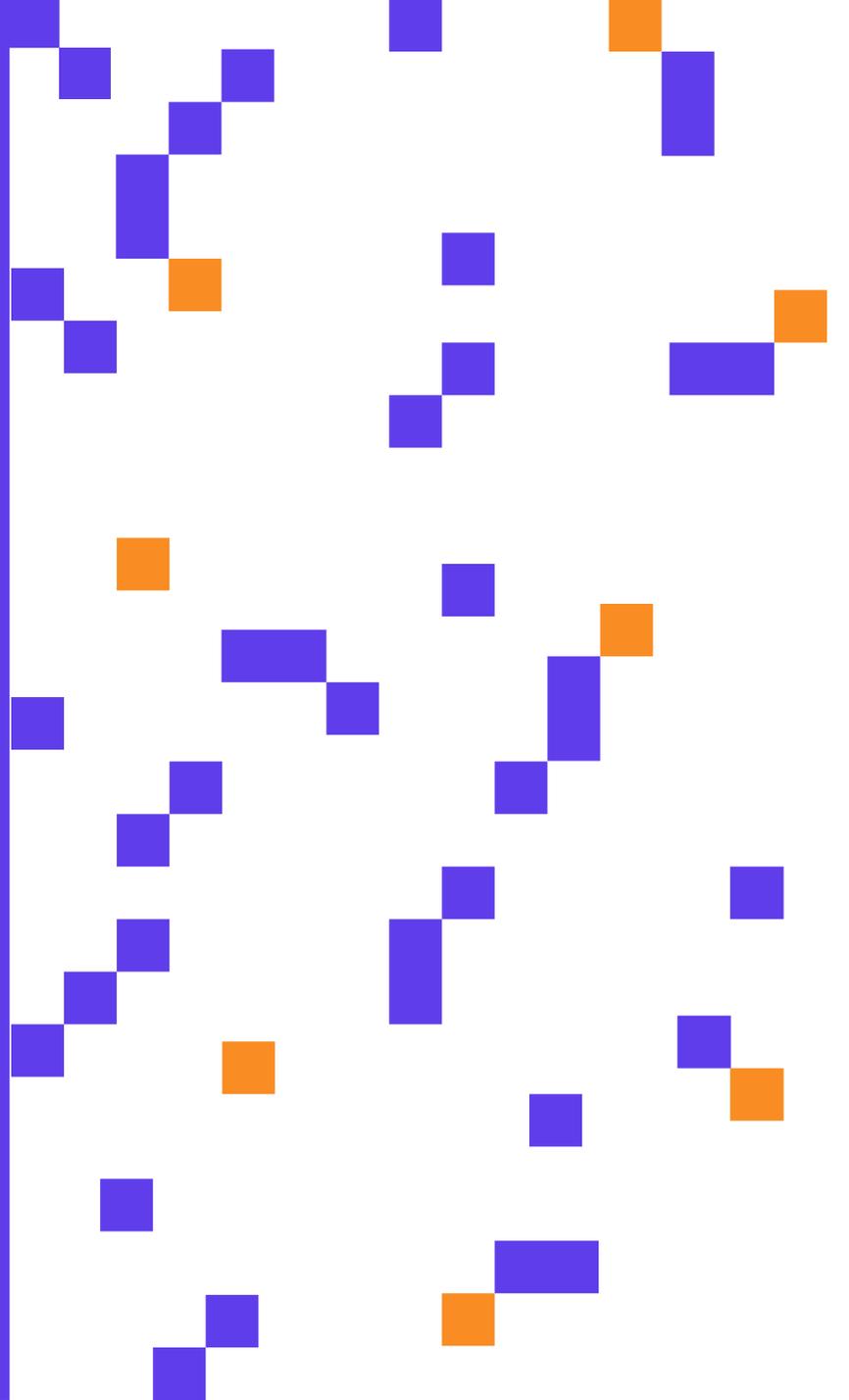
What to characterize?

- ❑ **Dimensions:** Standardized (ISO 25012) → e.g., accuracy, completeness, consistency, ...

How to characterize it?

- ❑ **Metrics:** Not standardized → leads to inconsistent between assessment frameworks
- ❑ **Need:**
 - ❑ A **standard mapping** of metrics to dimensions
 - ❑ Facilitates **design of data quality processes** across data spaces
 - ❑ Improves **interoperability** of quality assessments between users

Data Valuation



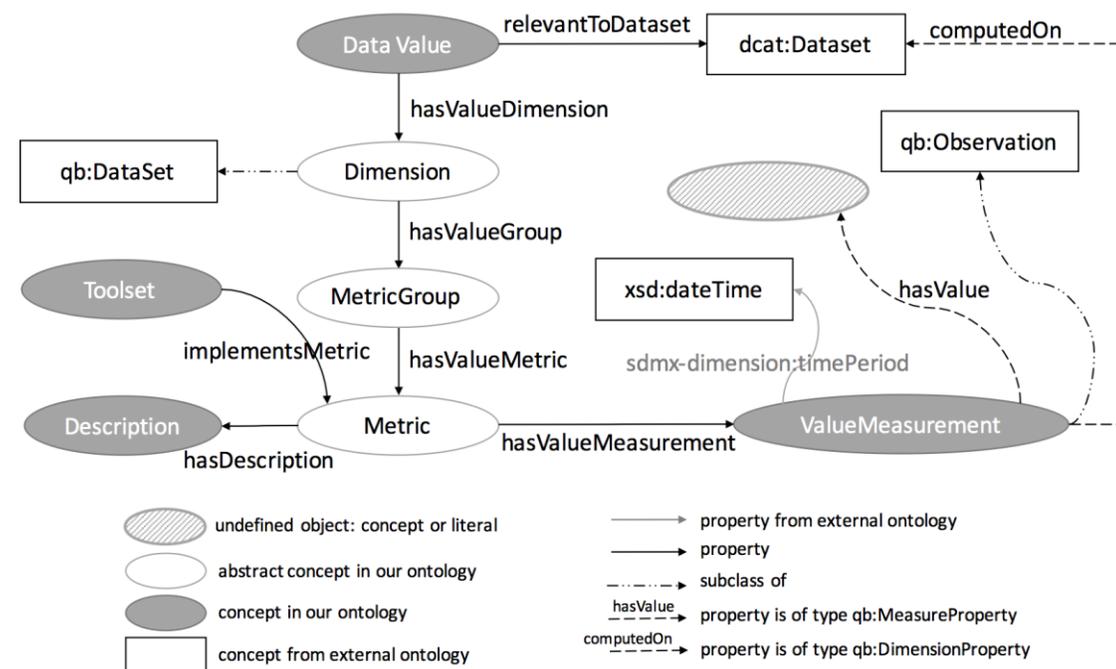
Data Value

- ❑ PISTIS – Dimensional Valuation Model [Fleckenstein et al., 2023]
 - ❑ Dimensions: Accessibility, Availability, Format, Age, Utility, Quality
 - ❑ Each of them can have several sub-dimensions
 - ❑ Availability: sharing, reselling, transfer, exclusivity
 - ❑ Age: maturity, recency, timeliness
 - ❑ Quality: completeness, validity, uniqueness, ...

DaVe – Data Value Vocabulary

[Attard and Brennan, 2019]

- Backbone concepts for expressing value: *Dimension, MetricGroup, Metric*
- Uses W3C Data Quality Vocabulary and DCAT
- Does not map directly to the PISTIS Valuation Methodology → alignment challenges
 - Difficult to express recursive composition of sub-dimensions into dimensions
 - Followed by re-aggregation of corresponding scores
- Currently not published, stored, nor maintained!



Source: [Attard and Brennan, 2019]

THANKS FOR YOUR ATTENTION.

Yury Glikman

Yury.Glikman@fokus.fraunhofer.de
<https://www.pistis-project.eu>



Funded by the
European Union

Quality Measurement and results

```
<https://www.pistis-project.eu/ns/voc#<distribution-uuid>/measure/<measurement-id>>
```

```
  a dqv:QualityMeasurement ;
```

```
  dqv:isMeasurementOf <https://www.pistis-project.eu/ns/voc#ExpectColumnValueLengthsToBeBetween> ;
```

```
  dqv:computeOn <https://develop.pistis-market.eu/catset/distribution/<distribution-uuid> ;
```

```
  dqv:hasQualityAnnotation <https://www.pistis-project.eu/ns/voc#<distribution-uuid>/measure/<measurement-id>/annotation> ;
```

```
  dqv:value "0.9"^^xsd:decimal .
```

Metric definition

```
<https://www.pistis-project.eu/ns/voc#ExpectColumnValueLengthsToBeBetween>
```

```
  a dqv:Metric ;
```

```
  skos:definition ""Expect the column entries to be strings with length between a minimum value and ..."" ;
```

```
  dqv:inDimension <https://www.pistis-project.eu/ns/voc#consistency> .
```

Quality Annotation with measurement parameters

```
<https://www.pistis-project.eu/ns/voc#<distribution-id>/measure/<measurement-id>/annotation>
```

```
  a dqv:QualityAnnotation ;
```

```
  oa:hasBody ""{"column": "4", "min_value": 2, "max_value": 5}"" ;
```

```
  oa:hasTarget <https://www.pistis-project.eu/ns/voc#<distribution-id>/measure/<measurement-id> .
```

Questions / Comments ?

Agenda

Insights from CEN/CENELEC JTC 25

Learnings from projects

Open discussion

Next steps

Next steps

Call to action...



Ideas

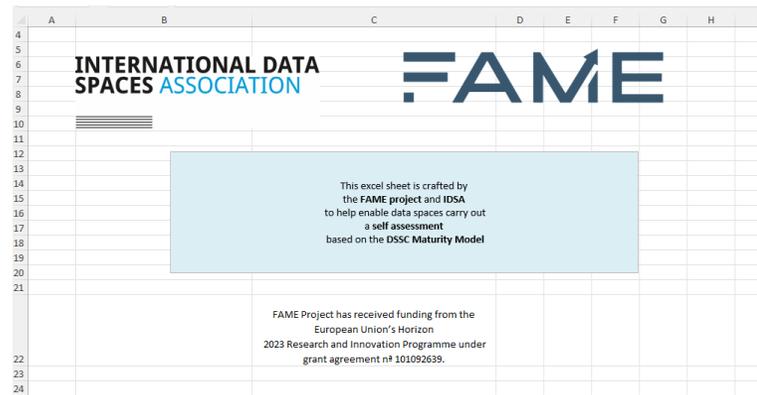
- Work together on **recommendations** for CEN/CENELEC JTC 25
- Collecting feedback on the DSSC Maturity Model, Data governance and Data quality
- A follow up **workshop** or **document** to share the conclusions from this workshop

Collecting feedback on the DSSC Maturity Model

Resources – will be circulated together with workshop slides

An Excel questionnaire

- Crafted by FAME and IDSA
- to help data spaces conduct a Self Assessment
- based on the DSSC maturity Model



A survey for providing feedback on the DSSC Maturity Model

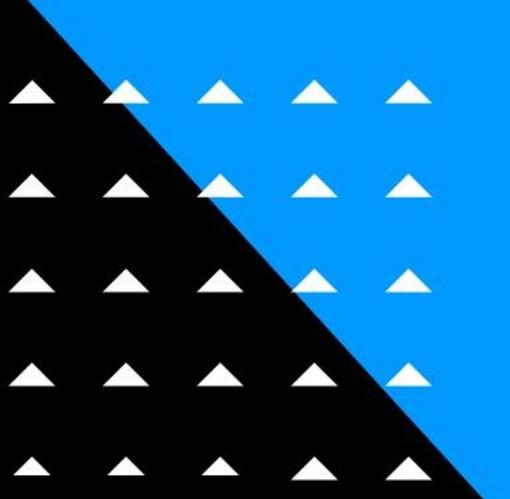
- MS Form: [Link](#)

The image shows a screenshot of a Microsoft Form. The form is titled "Your feedback on the DSSC Maturity Model" and is from the International Data Spaces Association. It includes a thank you message and a note about data collection. The form contains three required questions:

1. Your name *
2. Your surname *
3. Your email address *

Each question has a text input field with the placeholder "Enter your answer".





Thank you!



The logo for FAME consists of the letters "FAME" in a bold, blue, sans-serif font, with a stylized blue and orange icon to the left.



The logo for the International Data Spaces Association features the text "INTERNATIONAL DATA SPACES ASSOCIATION" in a bold, sans-serif font, with a stylized blue and orange icon to the left.