DATAWEEK

JOIN.LEARN.SHARE.GET VALUE

Navigating Challenges and strategies for data sharing and integration in the Data Space Era

CyclOps

DS2
DATA SPACE | DATA SHARE

CEDAR

NOUS Empowering Europe's Data Future

PLIADES

ORGANISED BY BDV BIG DATA VALUE ASSOCIATION

HOSTED BY COSMOTE

IN COLLABORATION WITH ATC, CORE|IC, FORTH, DEMOKRITOS, NEUROPUBLIC

In the workshop today…..

Data sharing and integration & leveraging Data Spaces

WHAT

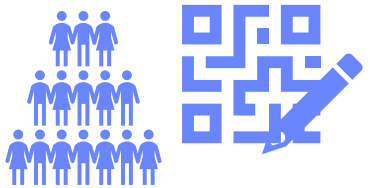
The user & their expected added value

Challenges and barriers

Solutions and Data Spaces alignment and support

HOW

Panellists present project insights



You participate via live forms

Discussion and generation of new insights

Panellists

MONICA CABALLERO

NTT DATA

EU R+D Project Manager
CyclOps project coordinator

JULIA PALMA

CeADAR, Ireland's Centre for AI

Programmes Manager
etami Co-Lead

RIINA LUOMA

VTT

DS2 Vice Coordinator

ELIAS DAKOS

Athens Technology Center

Project Manager
DS2 Project Partner

THEODOROS SEMERTZIDIS

CERTH

CEDAR Vice Coordinator

MORAITIS PANAGIOTIS

NetCompany

KOSMAS TSIAKAS

CERTH-ITI

Research Associate
PLIADES project

CyclOps

DS2
DATA SPACE | DATA SHARE

CEDAR

NOUS Empowering Europe's Data Future

PLIADES

DATAWEEK

JOIN.LEARN.SHARE.GET VALUE

DATA SHARING, INTEGRATION AND DATA SPACES-RELATED SOLUTIONS

PROFILING THE USER AND THEIR EXPECTED OUTCOMES

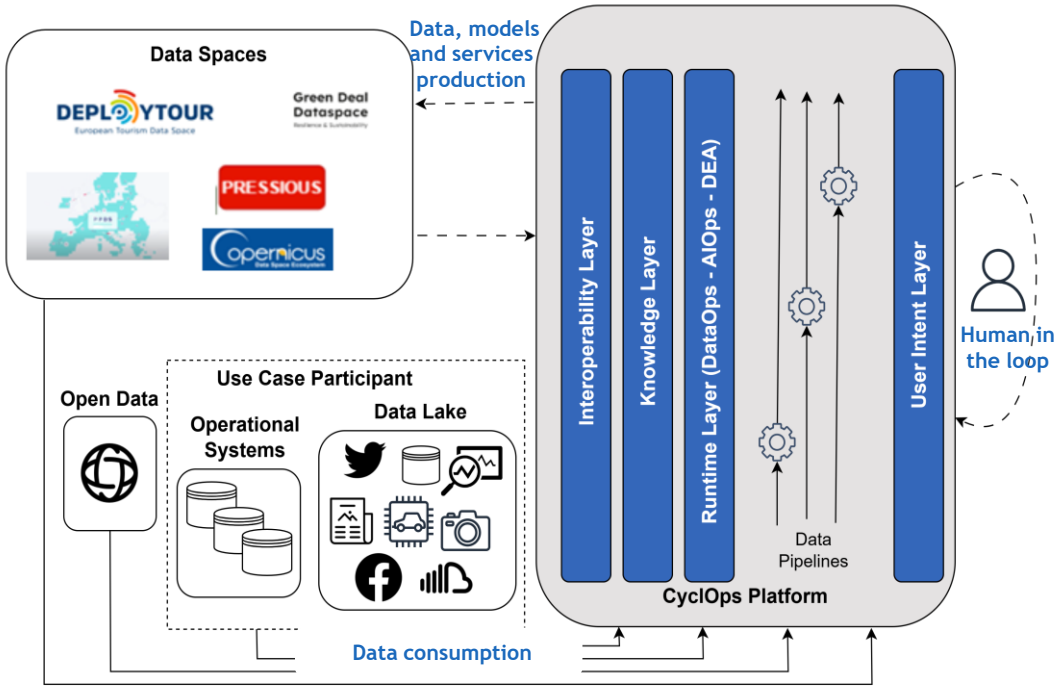
CyclOps

Automated end-to-end data life cycle management for FAIR data integration, processing and re-use

OBJECTIVE

Provide **interoperable** automatic management, governance, and maintenance of the entire data life cycle for large-scale volumes of data generated in heterogeneous distributed sources to **enable data sharing and exchange in data spaces and enable AI-based data-driven applications for all players, business and research alike.**

1. CyClops platform to automate generation of data processing pipelines
   - From the consumption of data from multiple sources to the serving of data, models, and services ready for data spaces
2. Deployment of 4 Use Cases linked to the potential future EU Data Spaces
   - Tourism and Mobility, Green Deal, Public Procurement, Manufacturing



36 months
January 2024 - December 2026

Budget - EU funding
8.894.072,50 €

24 partners from 11 countries

The user and its the expectations for CyclOps

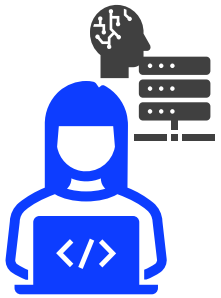
- CyclOps set of tools is targeted to any **organization that want to leverage, cross and merge their data with data from relevant dataspaces** to:
  - **Extract insights** - Research-oriented purposes
  - **Create value-added services** for their clients (business-alike purposes)

***TOURISM**: Create tools for DMOs to dynamically manage the demand based on DS data and dynamic sources*

***CLIMATE RISK**: Provide tools and re-usable data for enhanced climate crisis assessment and studies merging complex raw data*

***PUBLIC PROCUREMENT**: Create tools for policy makers to assess and predict the impact of policies*

***PRINTING IND:** Share data to create predictive maintenance AI-bases services*



Developers – Data scientists

Data Engineers, Data Analysts, Data Architects

Implement and manage data pipelines: take data from diverse sources and process it to generate specialized AI models or other insights

!!! Great diversity of skills maturity and levels (data management, AI, …).

Knowledge Scientists/ Domain experts

Configuration Support: manage the creation and handling of semantic assets, crucial for data integration and interoperability

Low availability in SMEs

Data Space, Data Sharing 2.0

The **DS2** project produces an approach for **inter data space interoperability** and **Intersector DataSpace Toolkit (IDT)**, which is deployable by the stakeholders of each Dataspace to enable cross space data sharing.

- Decision support for **complex data lifecycle sovereignty and governance.**
- **Human in the loop** modules addressing social and cultural differences.
- Framework building blocks for **data access and governance.**
- **Four use cases:** City Scape, Green Deal, Precision Agriculture, X-Intersector.

create business/ knowledge, etc.

The users and their expectations for

DS2 DATA SPACE | DATA SHARE

DS2 ideal users? Not an “ideal user profile ” per se, but more of a “universal user”. Anyone who wants to share or consume data via DS2 can participate; technical, non-technical, scientists, researchers, business, government and many more.

At a higher level: **DS participants** are the ideal users of DS2.

At a second level:

- **Data producers and Data consumers = prosumers (Interchangeable roles).**

- **DS facilitators/ brokers/ operators.**

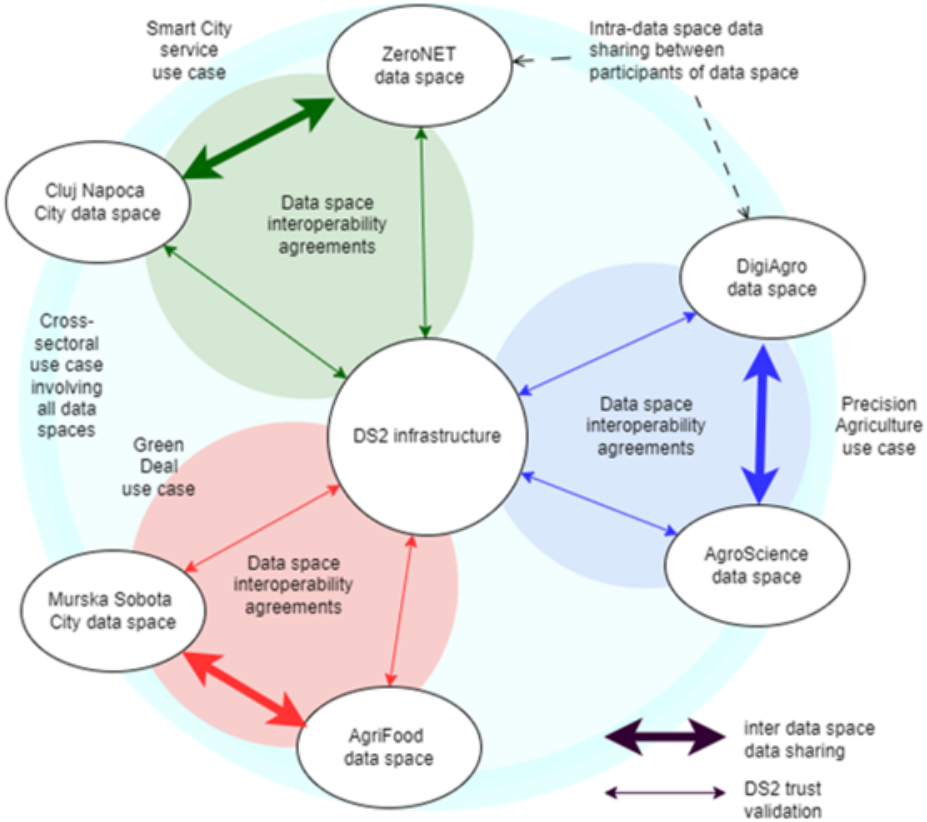
At a use case level, users ranging from data scientists, business owners to public administrators:

- **Use case 1, City Scape:** Universities, Municipalities, City departments, Third-party orgs, Start-ups.

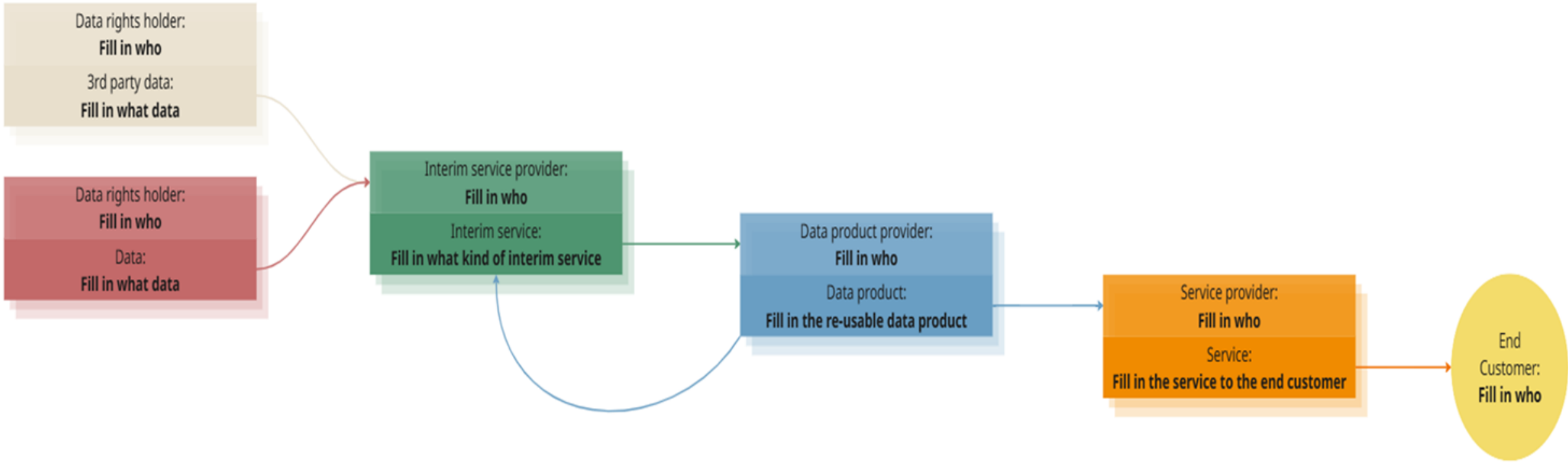
- **Use case 2, Green Deal:** Municipalities, Local business, Green economy.

- **Use case 3, Precision agriculture:** Agronomists, Farmers, Retailers, Supply Chain (Canning companies, Farmer suppliers), Meteorologists, Consultants for Data-sharing.

- **Use case 4, X-Intersector:** All the above!



User expectations: value creation process via data products



Common European Data Spaces and Robust AI for Transparent Public Governance

- CEDAR is developing an architecture that builds upon Common European Data Spaces and local data sources to identify corruption patterns in various scenarios
- CEDAR is bringing AI at the core of CEDS design patterns
- CEDAR will be validated in 3 different pilots and multiple scenarios
- CEDAR examines cross-border, cross-domain hook with different CEDS to exploit and share data beyond the obvious use

CEDAR AI components: LLMs, NLP, NLU, Computer Vision, Deep fakes, Satellite data, Audio analysis, Econometrics, Data fusion

Why interoperability

- Cross-organisational/ cross-border exchange of data will answer many questions about corruption and fraud
- Similar behavioural patterns exist
- Corrupted companies' traces may be identified in the most peculiar data sources

Challenges

- Abundance of data but no structure in a pan-European level
- CEDS are still of low maturity – CEDS connectors still under development
- Currently no CEDS available for external data consumers

Insights

- First, the digitalization of sources – still not horizontally available in European public sector
- Second, missing a mature, commonly agreed and validated technical solution

SELECTED PILOTS

- Transparent Management of National Recovery and Resilience Plans Funds in **Italy**
- Transparent Management of Public Healthcare Funds in **Slovenia**
- Transparent Management of Foreign Aid for Rebuilding **Ukraine**

CEDAR Users and Expectations

IT Italy - Transparent Management of NRRP Funds

Users:

- Regional Government Directors (e.g., Region Friuli Venezia Giulia)
- Regional Departments Employees (e.g., Department of Infrastructure and Transport)
- Public Buyers (e.g., Procurement Manager of Municipality of Trieste)

Expectations:
Greater transparency and streamlined oversight of NRRP spending across regional bodies and procurement offices.

SI Slovenia – Transparent Management of Public Healthcare Funds

Users:

- Police officers
- Ministry of Health auditors
- IT dept. of Ministry of Digital Transformation
- Procurement officers of Hospitals and Healthcare Providers (e.g., SBC, general and specialist hospitals, health centers)
- Suppliers and Bidders

Expectations:
An automated, transparent procurement platform that ensures fair competition and efficient allocation of healthcare funds.

UA Ukraine - Transparent Management of Foreign Aid for Reconstruction

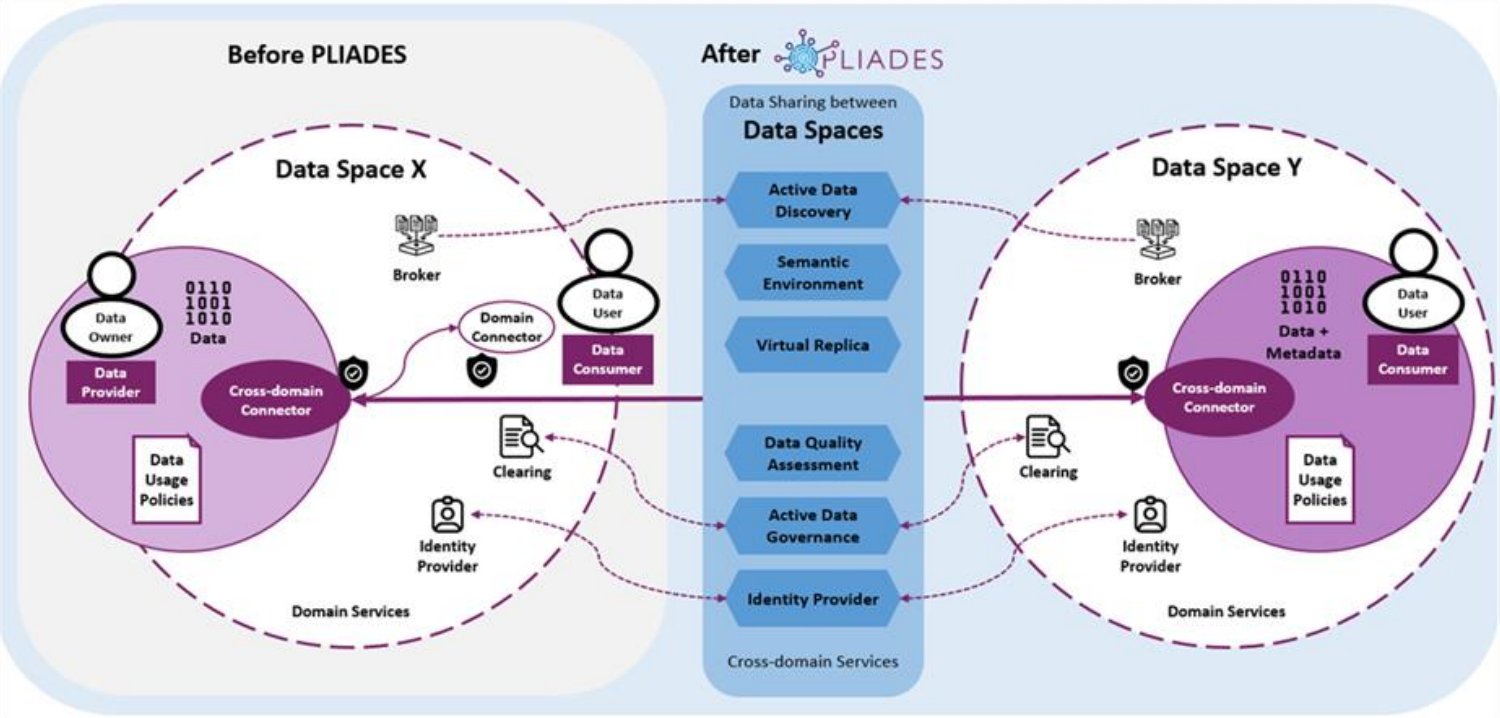
Users:

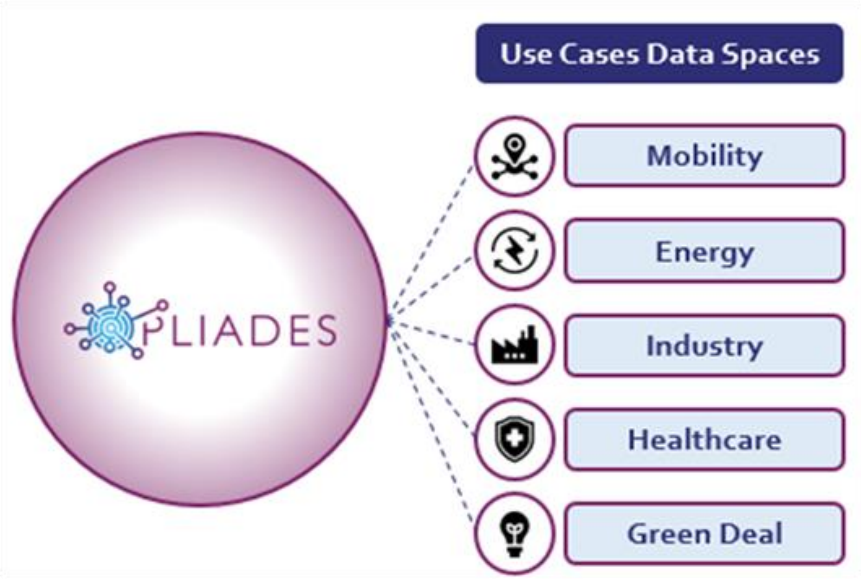
- State Agency for Restoration and Infrastructure Development (ARU) officers examining offers and tenders
- Anti-Corruption and Law-Enforcement officers (e.g., NABU, NACP, National Police, SBU, ESBU) identifying potential corruption links

Expectations:
Automated fraud detection and procurement monitoring to safeguard reconstruction funds and support anti-corruption efforts.

AI-Enabled Data Lifecycles Optimization & Data Spaces Integration for Increases Efficiency & Interoperability

PLIADES establishes a secure & seamless **Data Exchange Ecosystem** by integrating **Diverse Data Spaces**





Funded by the European Union

https://www.pliades-project.eu/

42 months
January 2024 - June 2027

Budget - EU funding
8.999.820€

27 partners from
11 countries

DATAWEEK
JOIN.LEARN.SHARE.GET VALUE

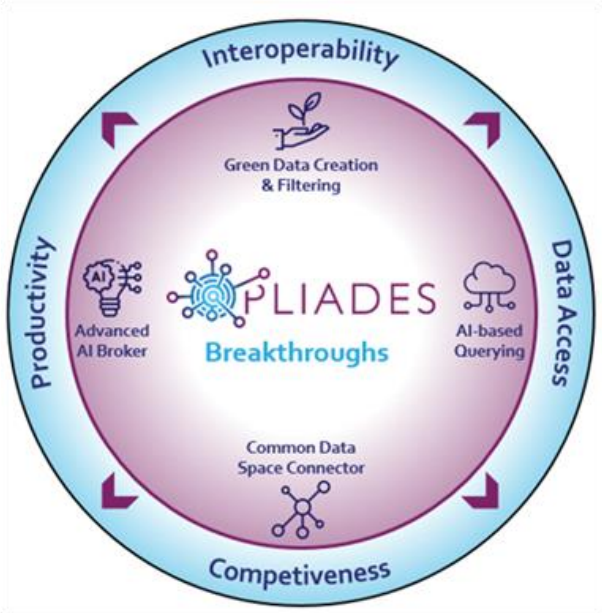
The user and its expectations for PLIADES

Users

- Data Engineers, Data Scientists, AI Developers
- IT Departments, System Integrators
- Researchers, Domain Experts, Engineers
- Technical Managers, R&D Leads

Needs:

Tools and infrastructure for secure, seamless, and explainable AI-driven data management across complex environments



Expectations for PLIADES:

Enhancing Lifecycle Optimisation & Data Spaces Integration to:

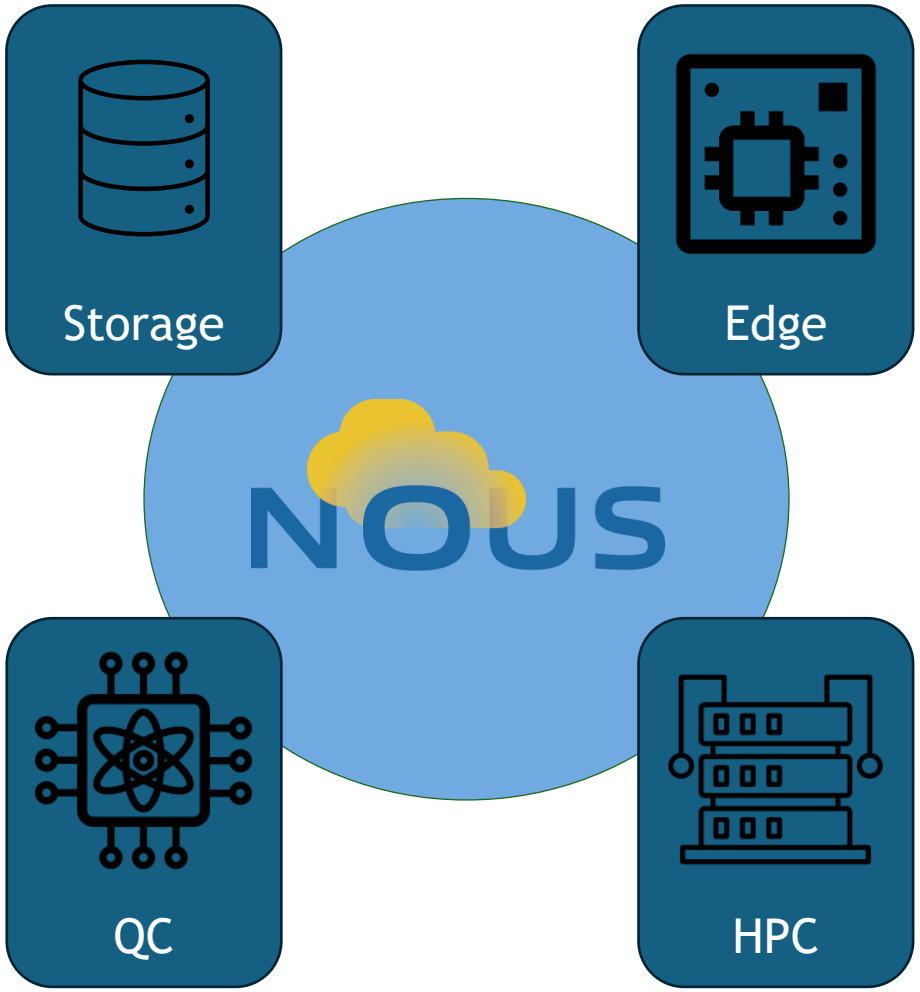
- Interconnect Diverse Data Spaces → Enhancing Efficiency & Interoperability
- Expand Data Lifecycles → Supporting Green Data Creation, Ownership, Reuse & Disposal
- Enable AI-driven Brokers → Facilitating Data Discovery & Integration
- Provide Data & Services → Improving *CCAM*[1] & *ADAS/AD*[2] & *HRI*[3]

NOUS

Empowering Europe's Data Future

NOUS will develop the architecture of a European Cloud Service in the form of Infrastructure-as-a-Service (IaaS) /Platform-as-a-Service (PaaS) cloud provider based on the **Simpl Open** Architecture that will allow **computational** and **data storage** resources to be used from:

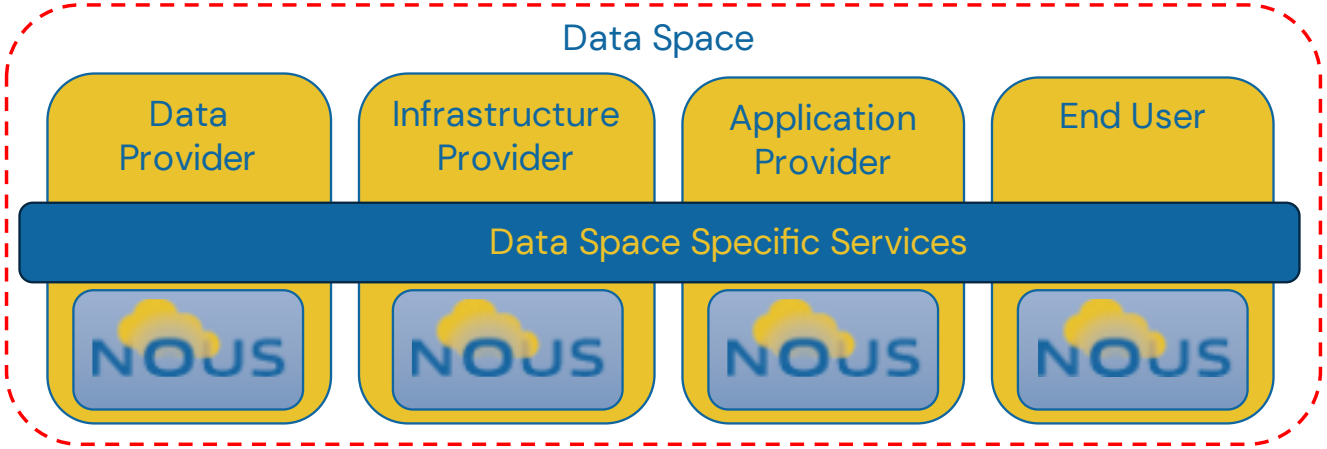
- **Edge devices**
- **Supercomputers** through the HPC network
- **Quantum Computers**



The user and its expectations for NOUS

Empowering Europe's Data Future

1. **Data providers** hold valuable data assets that they share to the NOUS ecosystem
2. **Infrastructure** providers allow users to run applications on their infrastructure
3. **Application** providers share applications and algorithms with end users
4. **End users** consume assets and resources



Use Case #1
Perception of Connected Vehicles using camera data

Use Case #2
Energy Prediction and Energy Data Lifecycle Management

Use Case #3
Crisis Management and Civil Protection Platform

Use case #4
Scientific Data HPC Storage and AI Analytics

QUESTIONS FOR THE AUDIENCE

Reply to the questions online

- Select the most suitable combination of the role of your organization within the data ecosystem and your profile within your organization
  - Data owner - data source / Technical - Data Scientist, Engineer, Developer, Architect
  - Data owner - data source / Technical - Knowledge scientist
  - Data owner - data source / Business (e.g. Consultant, Product Manager, Innovation lead, etc.)
  - Data solution provider / Technical - Data Scientist, Engineer, Developer, Architect
  - Data solution provider / Technical - Knowledge scientist
  - Data solution provider / Business (e.g. Consultant, Product Manager, Innovation lead, etc.)
  - End-user of data-based application / Technical - Data Scientist, Engineer, Developer, Architect
  - End-user of data-based application / Technical - Knowledge scientist
  - End-user of data-based application / Business (e.g. Consultant, Product Manager, Innovation lead, etc.)
  - Other
- What value do you expect to obtain from data sharing, data integration and Data Spaces?

Select the
within the

Scan the QR or use link to join

https://forms.office.com/e/zUZCdfgGEC

Copy link

Data owr
/ Tec
Scien
Develc

Data solution provider / Technical - Data Scientist, Engineer, Developer, Architech

Data solution provider / Technical - Knowledge scientist

Data solution provider / Business (e.g. Consultant, Product Manager, Innovation lead, etc.)

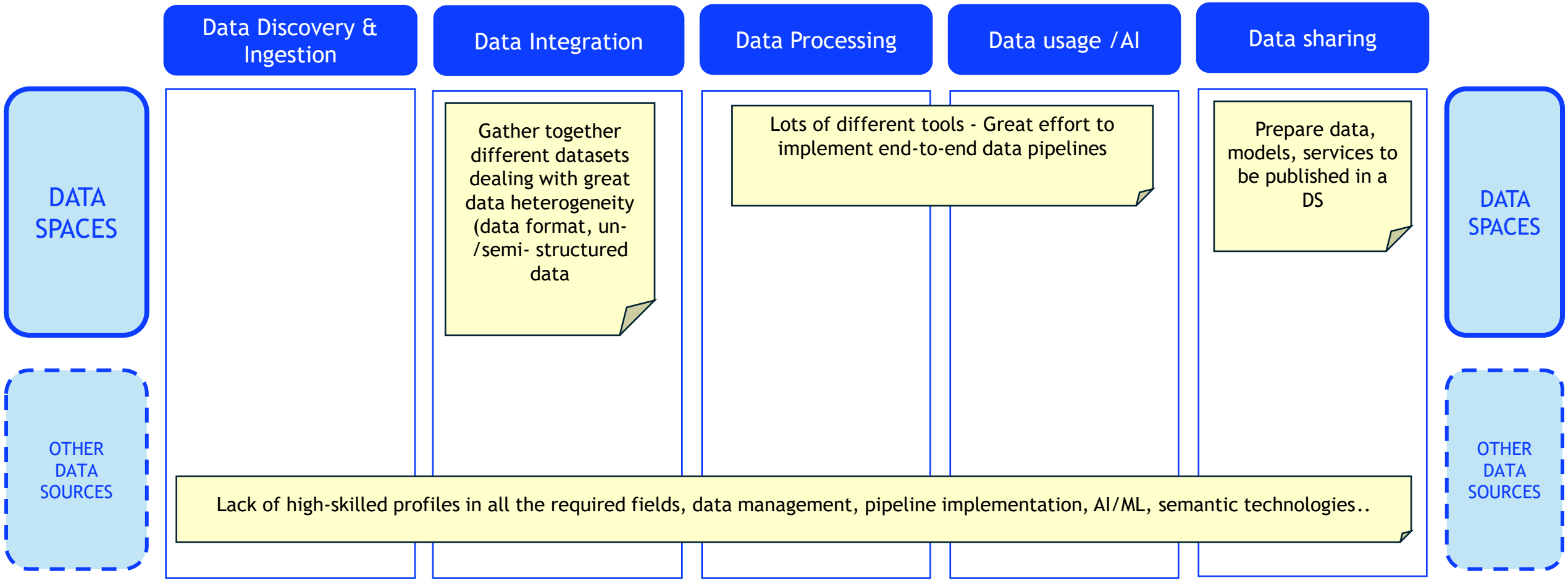
DATAWEEK

JOIN.LEARN.SHARE.GET VALUE

WHAT IS CURRENTLY PREVENTING AND/OR SLOWING DOWN NEW DATA-BASED SERVICES?

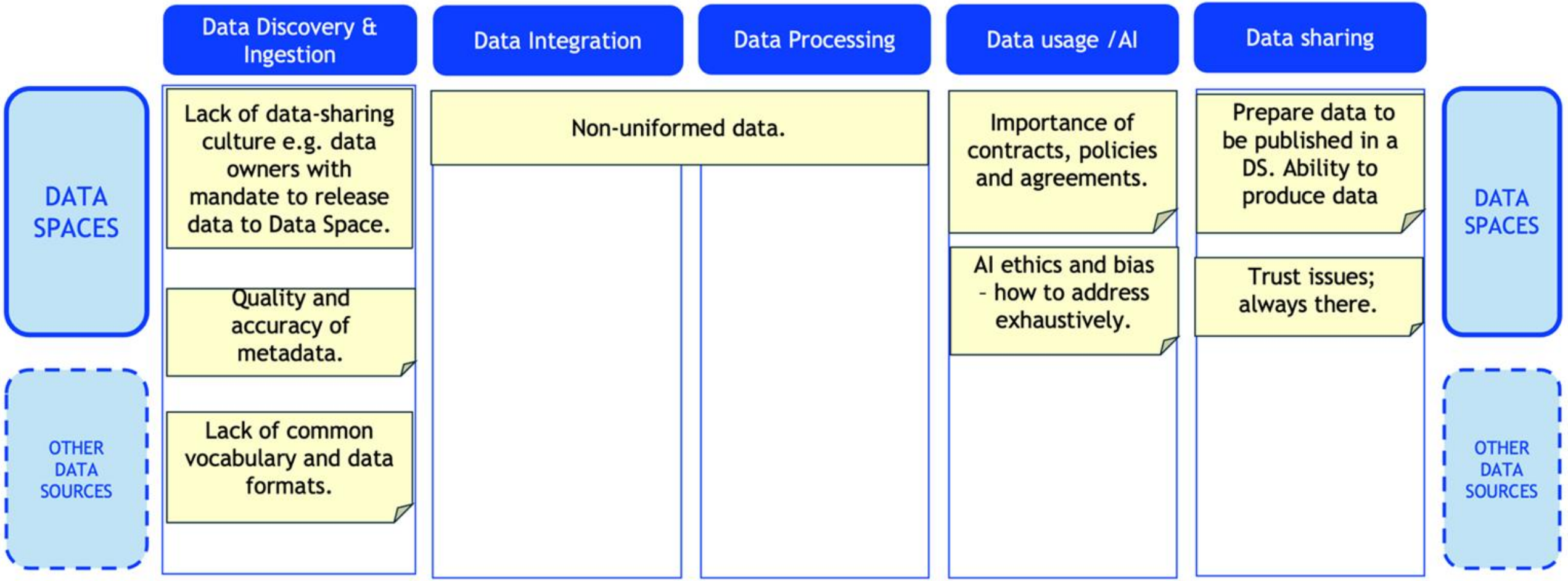
CHALLENGES AND BARRIERS

Pain points identified in CyclOps



Pain points identified in DS2

DS2
DATA SPACE | DATA SHARE



Pain points identified in CEDAR

DATA SPACES

OTHER DATA SOURCES

Data Discovery & Ingestion	Data Integration	Data Processing	Data usage /AI	Data sharing
It is not possible to find data from different domain data spaces	no consolidated integration procedure and architectural blocks	Many available tools - very few tested and validated for the purpose.	Data quality is not guaranteed	Privacy and data regulatory compliance
Different digital transformation maturity levels of Member states - digitization of public domain data is relatively low	Data models homogenisation needed		No common way of evaluating the performance of AI components for different purposes	lack of trust for publishing data related to public governance
			No clear/easy roadmap to follow AI Act compliance on AI modules	

Check our pre-standardisation document in https://cedar-heu-project.eu/

DATA SPACES

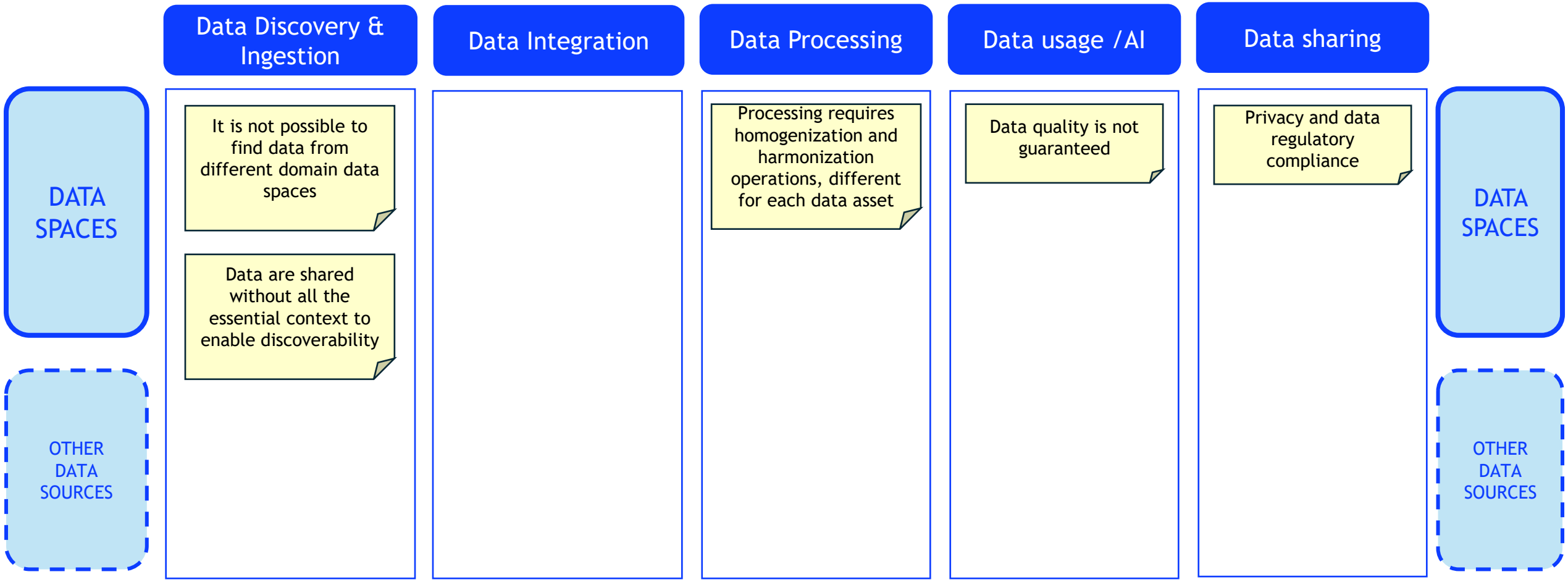
OTHER DATA SOURCES

Pain points identified in NOUS

Empowering Europe's Data Future

	Data Discovery & Ingestion	Data Integration	Data Processing	Data usage /AI	Data sharing	
DATA SPACES	Existing and Mature Data Spaces is a scarcity	Dataspaces do not centralize data, which means data consumers must interact with many heterogeneous systems		FL is complicated to be implemented	Joining a data space requires **identity verification** and **governance compliance** Data space connectors require different than business as usual expertise DS to DS communication and data exchange is still complicated	DATA SPACES
OTHER DATA SOURCES						OTHER DATA SOURCES

Pain points identified in PLIADES



QUESTIONS FOR THE AUDIENCE II

Reply to the questions online

- For you/your client (the end-user you want to represent), which phase/area typically causes the most bottlenecks?

  (Rank in order of impact: 1 – Highest impact, 4 – Lowest)

  - Discovery & Ingestion
  - Data Integration
  - Data Processing
  - Data Usage / AI
  - Data Sharing

- Indicate the 3 most relevant pain points. Identify any other relevant area not considered.

Main pain points for data sharing, integration and leveraging Data Spaces

27 may 2025

DATAWEEK25 - Navigating Challenges and strategies for data sharing and integration in the Data Space Era

Scan the QR or use link to join

https://forms.office.com/e/Q1fMfCm66T

Copy link

Select the most su
within the data ec

Data owner - data so
/ Technical - Data
Scientist, Enginee
Developer, Archite

Data solution provider / Technical - Data Scientist, Engineer, Developer, Architech

Data solution provider / Technical - Knowledge scientist

Data solution provider / Business (e.g. Consultant, Product Manager, Innovation lead, etc.)

DATAWEEK

JOIN.LEARN.SHARE.GET VALUE

SOLUTIONS

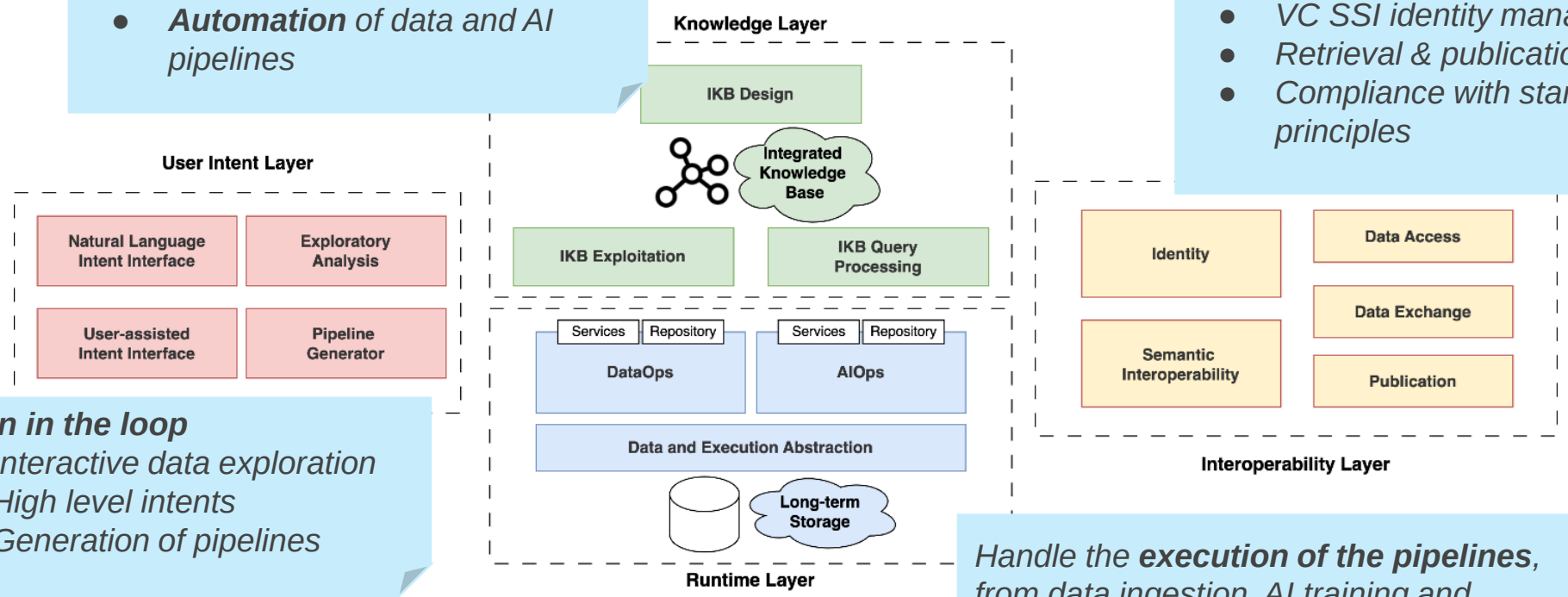
ADDRESSING MAIN PAIN POINTS & PROMOTING AND SUPPORTING DATA SPACES

The approach of CyclOps

- ***Homogeneous view of data*** *sources (semantic technologies)*
- ***Automation*** *of data and AI pipelines*

Bridge with Data Spaces *and other external sources*

- *VC SSI identity management*
- *Retrieval & publication of data*
- *Compliance with standards and FAIR principles*



Human in the loop

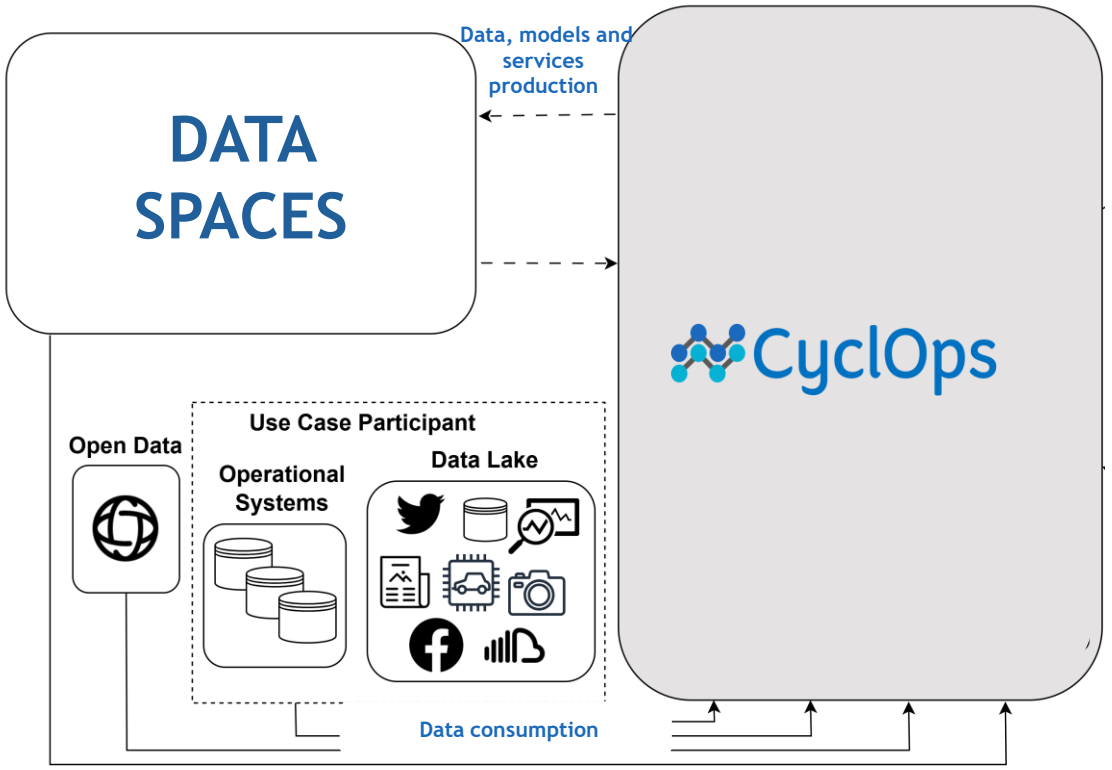
- *Interactive data exploration*
- *High level intents*
- *Generation of pipelines*

Handle the ***execution of the pipelines****, from data ingestion, AI training and deployment in distributed environments.*

The approach of CyclOps

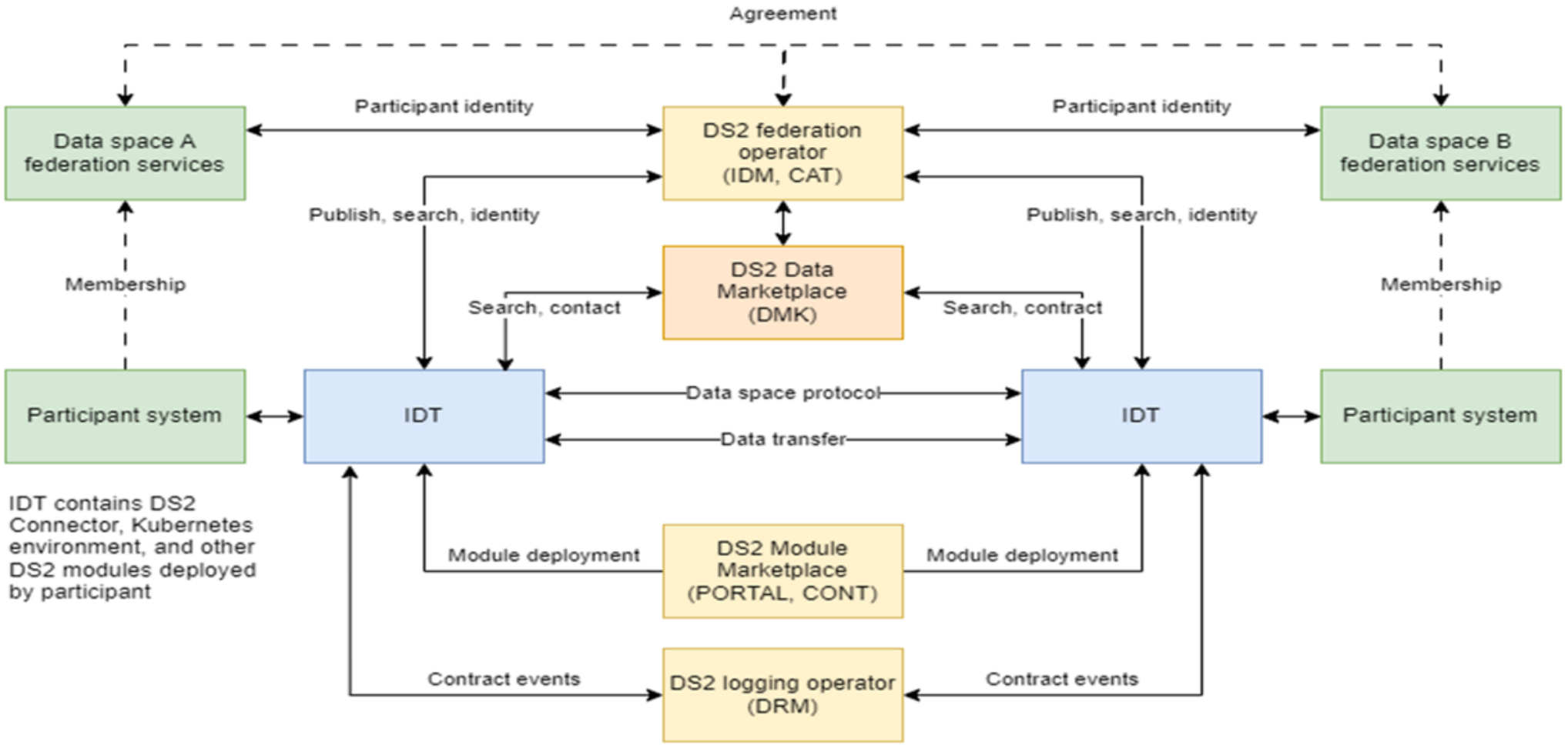
Supporting and promoting Data Spaces

- Facilitating the access and consumption of data from Data Spaces to a wide range of organizations profiles, enabling the exploitation of data and the integration with other external data sources for a given UC.
- Promoting population of DataSpaces, through specific services that prepare data, services and models to be published in Data Spaces.
- Key functionalities **do not overlap with DSSC building blocks**
  - The Interoperability layer brings DSSC components (VC, DCAT for catalgoues, etc..)
  - All components follow standards and recommendations of DSSC to maximize interoperability.



The approach of DS2

DS2
DATA SPACE | DATA SHARE



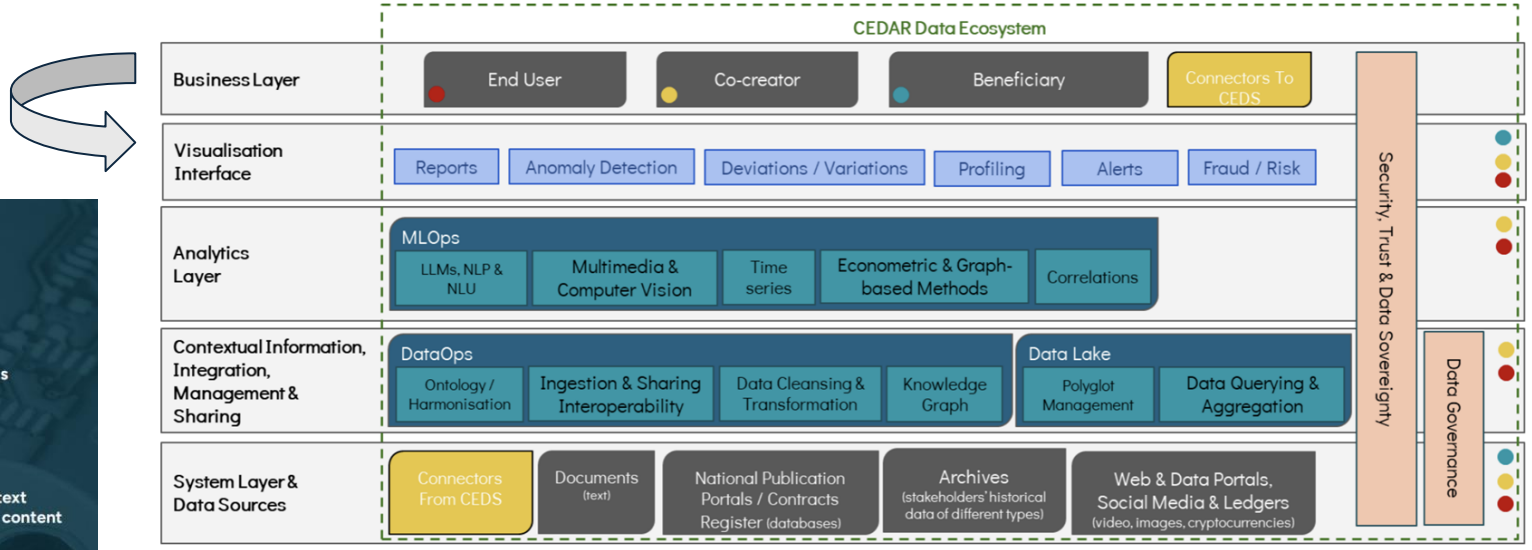
The approach of

DS2 is a software development project consisting of:

- module run-time environment with DS2 connector,
- module marketplace that facilitates deployments and where user can select modules case-by-case basis,
- set of modules implements which enable data functionalities between dataspaces, and
- connection capabilities with other existing dataspace marketplaces.

The approach of CEDAR

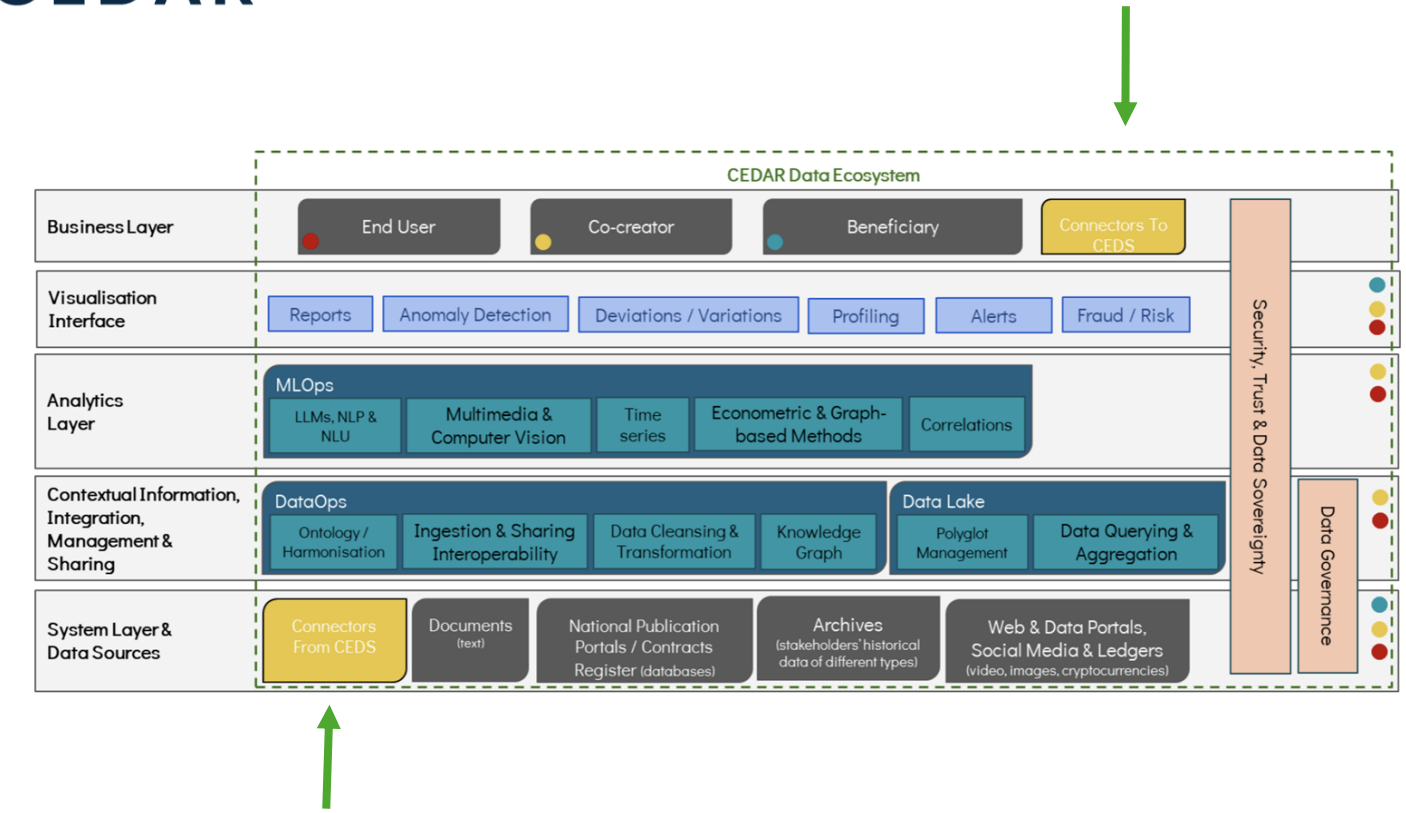
- Multiple AI tools
- A simple AI/CEDS integration architecture



- easy to follow AI/CEDS pipelines: seamless integration of CEDS, Processing and Analysis Pipelines and Operational UIs

The approach of CEDAR

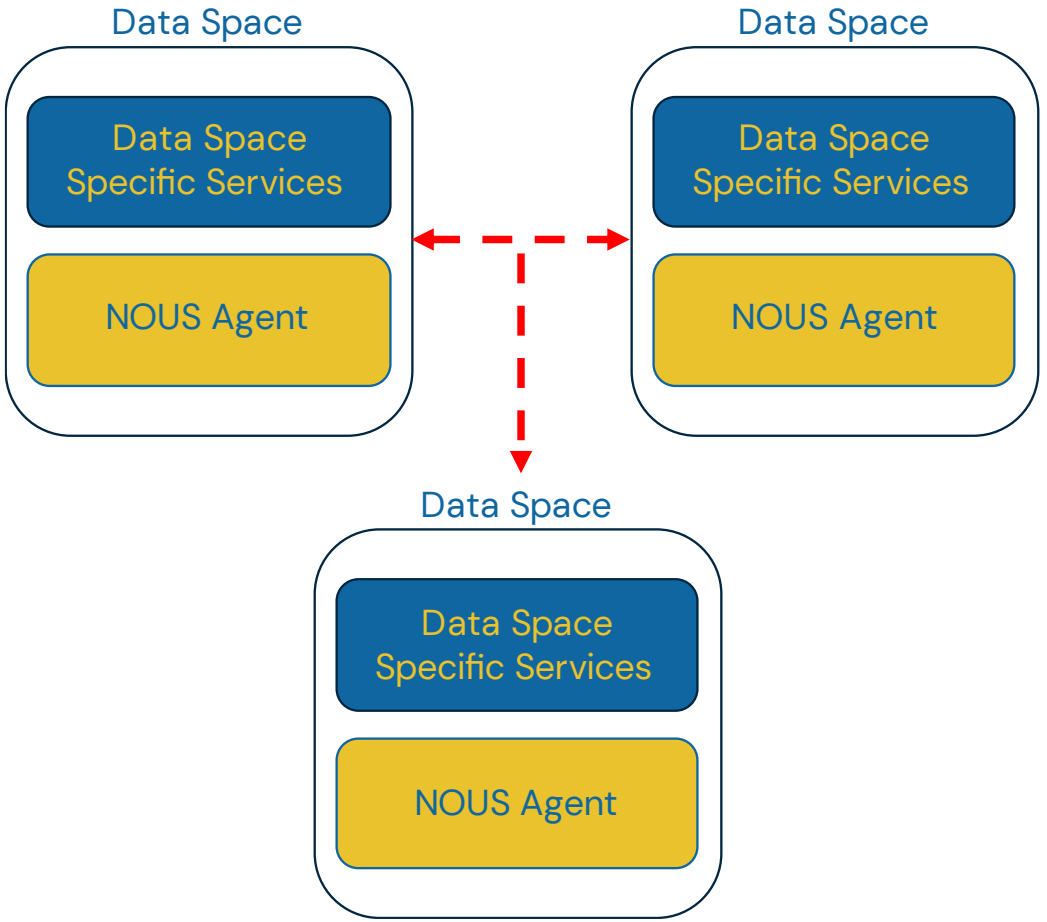
- Provide connectors to and from CEDS to enrich the data in the analysis procedure and share national datasets for cross border governance scenarios.
- Make the CEDS data sharing transparent to the end-user with integration in the operational UIs and pipelines.
- Using Sovity repository to build upon our integrated solution.



The approach of NOUS

Empowering Europe's Data Future

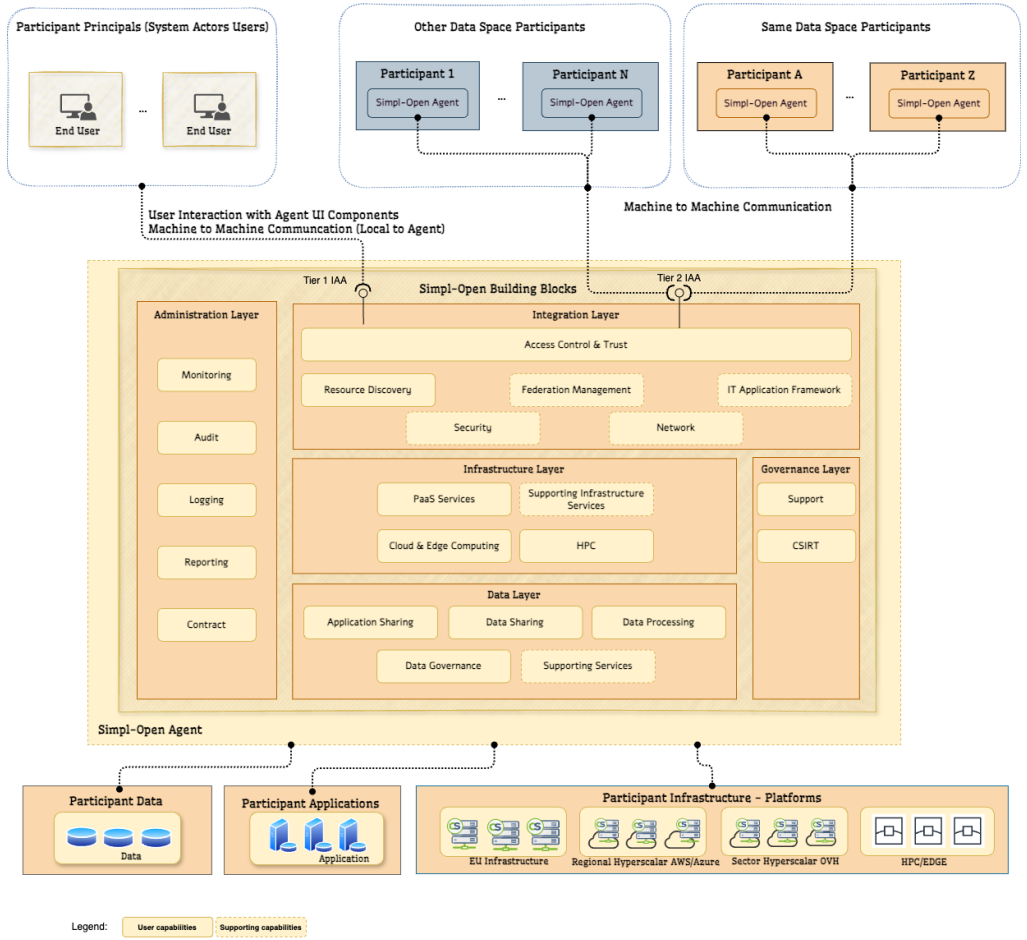
- Middleware Solution to federate data, software and infrastructure.
- Abstract component that actors need to deploy to become part of the ecosystem
- Deployment of the middleware in different "flavours" depending on the type of participant (actor)
- NOUS aims to keep the functionality at the edge to the fullest extend to reflect the decentralised purpose of a data space
- Centralised services are given minimal responsibility . Three types of centralised components were identified to support the distributed functionalities:
  - data, application and infrastructure catalogues
  - vocabulary providers
  - identification authorities



The approach of NOUS

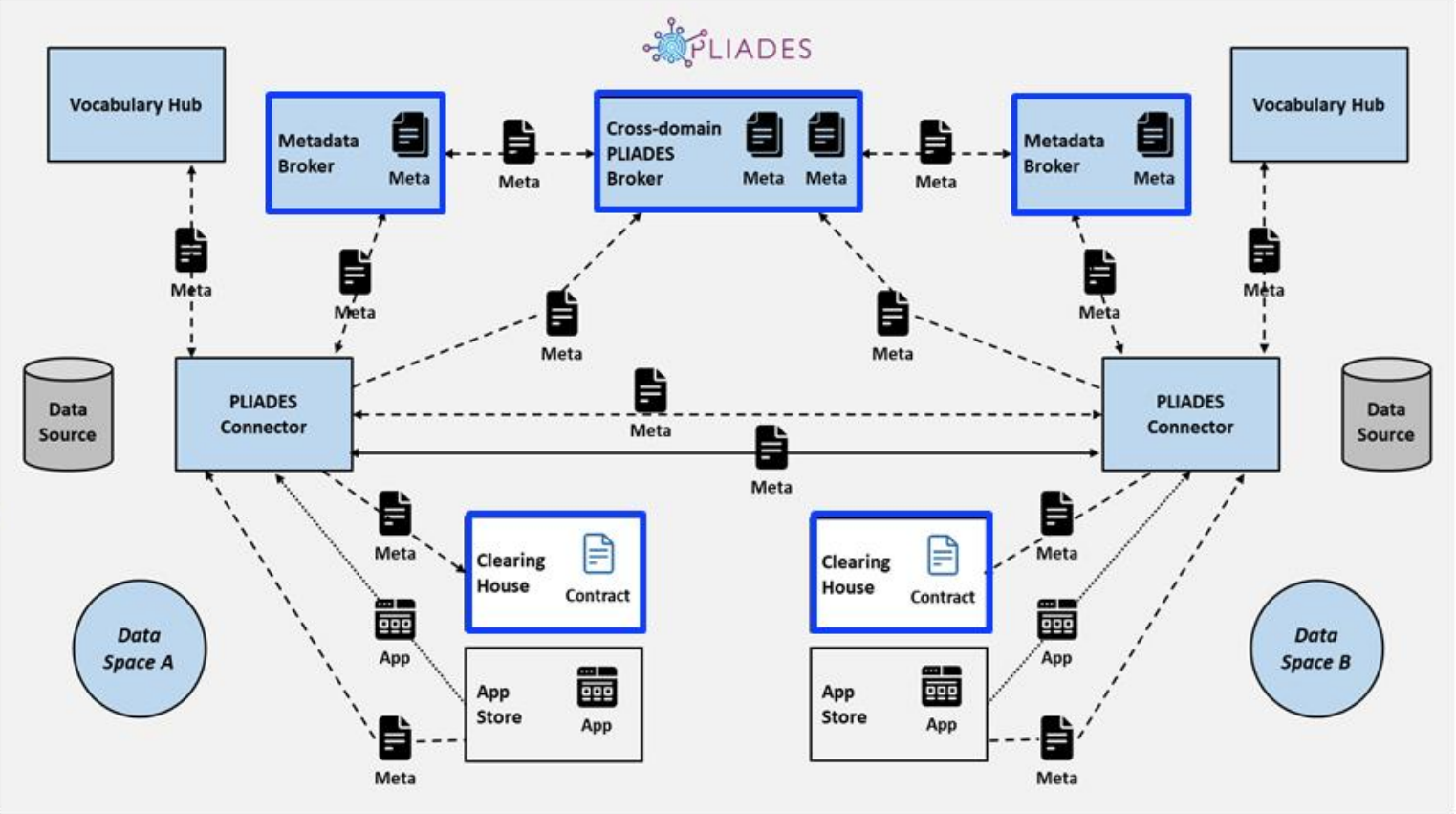
Empowering Europe's Data Future

- Build on the **Simpl-Open** Architecture that facilitates the creation of ecosystems for data and application sharing, as well as cloud infrastructure services.
- **Simpl Open** provides the basic blueprint to federate data, applications and infrastructure.
- Provides standardization in components and processes
- Identify relevant Business Processes and Functional Requirements and support the development of the open-source components



The approach of PLIADES

- Building upon the IDSA Reference Architecture Model layers, extending for cross-domain interoperability
- Use guidelines from Data Spaces Support Centre (DSSC) Blueprint building blocks complementary with IDS-RAM to address:
  - Business
  - Governance
  - Legal aspects

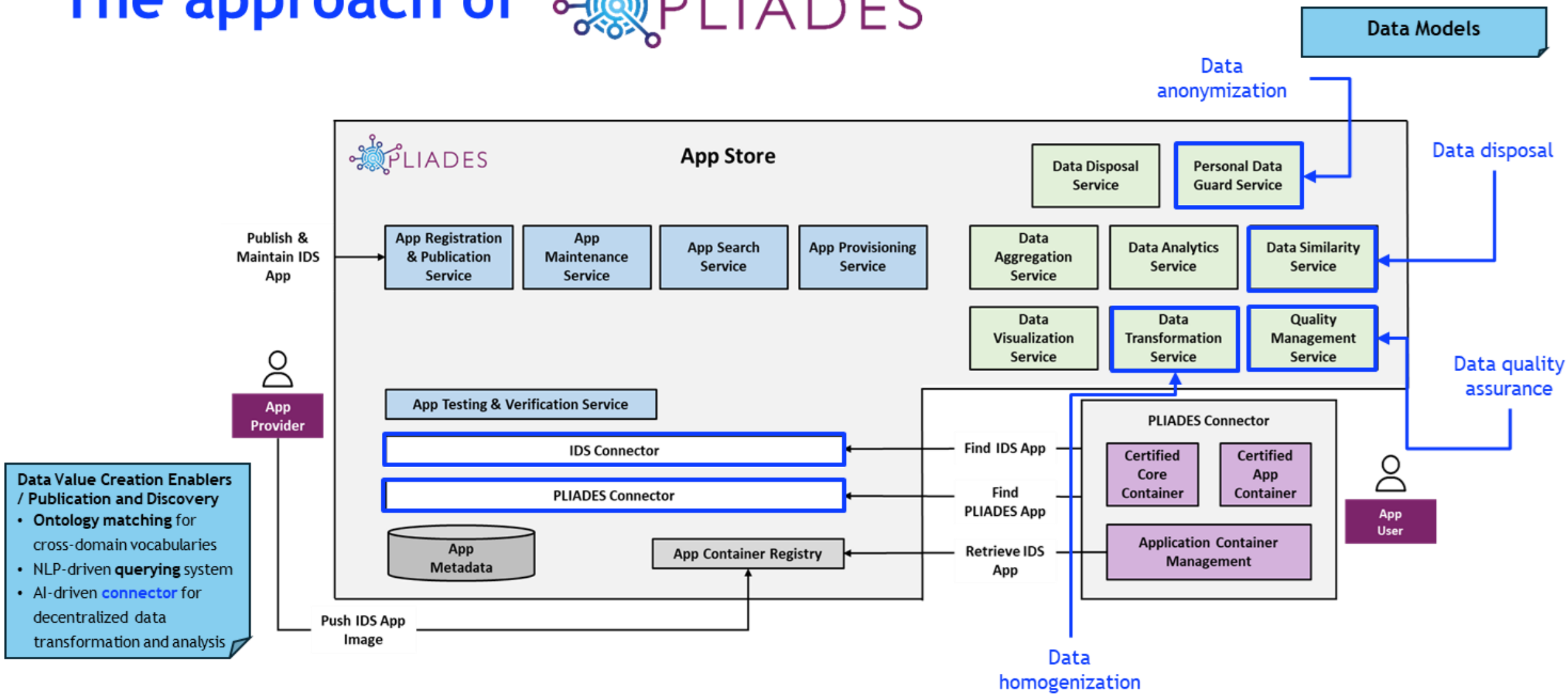


Data Value Creation Enablers / Publication and Discovery
- AI-driven Metadata Broker
- Cross-domain discoverability
- Recommendation engines & ranking system

Data Interoperability / Provenance & Traceability
- Hyperledger Fabric Blockchain (HFB) for the Logging service of a Clearing House
- Immutability, confidentiality, sovereignty, authenticity, integrity, traceability, provenance, non-repudiation
- Existing IDSA Clearing House does not cover the data provenance needs
- A trusted ecosystem for data providers and consumers with data-centric usage control

**There is no 1:1 matching between DSSC building blocks and PLIADES software components!*

The approach of PLIADES

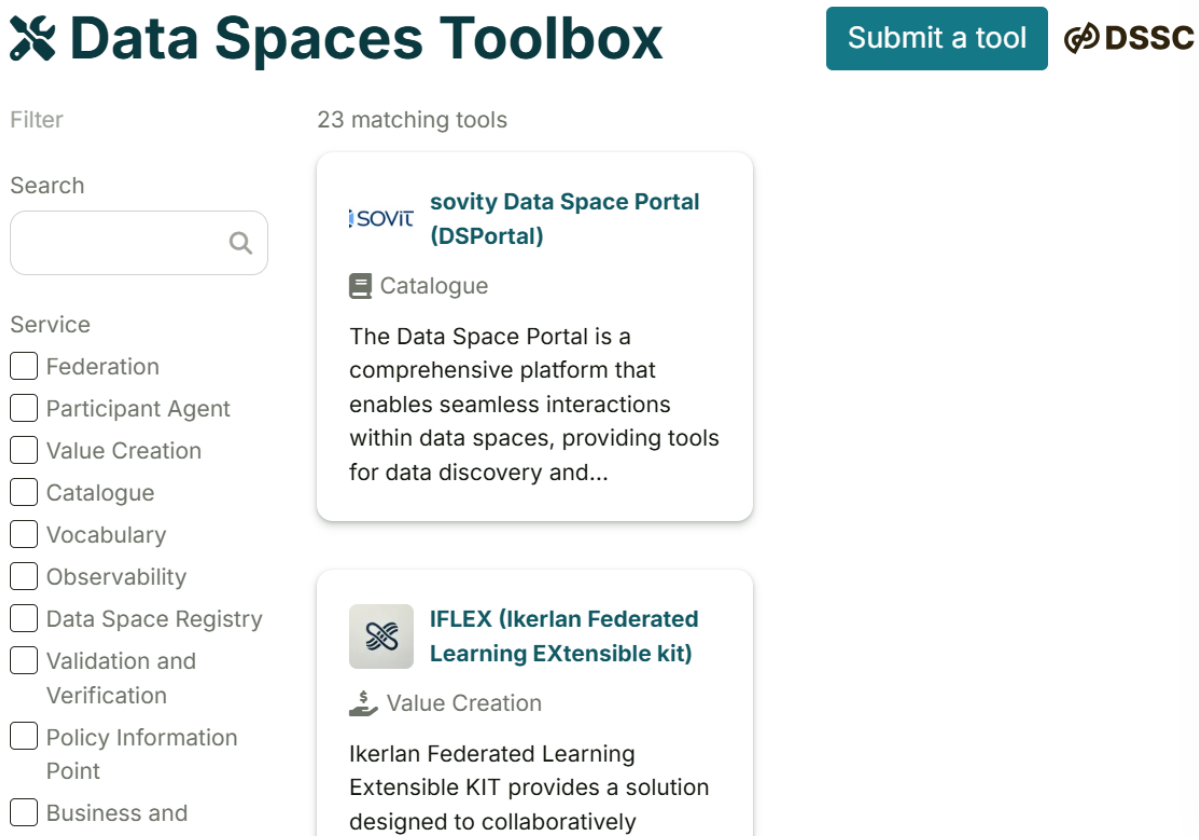


Questions to the DSSC for discussion

- How is the DSSC facilitating standardization or tool development that supports seamless data lifecycle management to access different Data Spaces?
- What tools, guidelines, or frameworks is the DSSC prioritizing to support key data lifecycle stages (e.g., ingestion, linkage, connectors)?
- How can the ongoing projects provide input or contribute to the evolution of DSSC's technical architecture, tools, or reference documents or implementations?

DSSC building blocks and toolbox

- How is the DSSC facilitating standardization or tool development that supports seamless data lifecycle management to access different Data Spaces?
- What tools, guidelines, or frameworks is the DSSC prioritizing to support key data lifecycle stages (e.g., ingestion, linkage, connectors)?
- How can the ongoing projects provide input or contribute to the evolution of DSSC's technical architecture, tools, or reference documents or implementations?



Thank you

CyclOps

DS2
DATA SPACE | DATA SHARE

CEDAR

NOUS
Empowering
Europe's Data
Future

PLIADES

Funded by
the European Union

DATAWEEK
JOIN.LEARN.SHARE.GET VALUE

ORGANISED BY
BDV BIG DATA VALUE ASSOCIATION

HOSTED BY COSMOTE

IN COLLABORATION WITH

ATC

CORE|IC

FORTH

DEMOKRITOS

NEUROPUBLIC