CEADS
COMMON EUROPEAN AGRICULTURAL DATA SPACE

MISTEA
Mathématiques, Informatique et Statistique pour l'Environnement et l'Agronomie

Bringing Semantics into the Common European Agricultural Data Space (CEADS) with AgroPortal

Dr. Filipi M. SOARES

18 May 2026

Co-funded by the European Union

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or European Commission. Neither the European Union nor the granting authority can be held responsible for them.

CEADS
COMMON EUROPEAN AGRICULTURAL DATA SPACE

MISTEA
Mathématiques, Informatique et Statistique pour l'Environnement et l'Agronomie

Agenda

Overview of today's presentation

1. A brief history of data spaces
2. What is a data space?
3. Key players in the data-space ecosystem
4. Governance, rules, and regulations
5. The Common European Agricultural Data Space (CEADS)
6. CEADS Use Case 2: Research and Innovation Data for Improved AI Algorithms Performance

Co-funded by the European Union

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or European Commission. Neither the European Union nor the granting authority can be held responsible for them.

A brief history of data spaces

Early 2000s

The idea of data spaces started as a research concept before becoming a European policy and infrastructure agenda

The concept of **dataspaces** first emerged in the data-management research community in the mid-2000s.

A key starting point is the 2005 paper by **Franklin, Halevy, and Maier**, *From Databases to Dataspaces*, which proposed dataspaces as a new abstraction for managing heterogeneous and distributed data sources.

From Databases to Dataspaces:
A New Abstraction for Information Management

Michael Franklin
University of California, Berkeley

Alon Halevy
Google Inc. and U. Washington

David Maier
Portland State University

Abstract

The development of relational database management systems served to focus the data management community for decades, with spectacular results. In recent years, however, the rapidly-expanding demands of "data everywhere" have led to a field comprised of interesting and productive efforts, but without a central focus or coordinated agenda. The most acute information management challenges today stem from organizations (e.g., enterprises, government agencies, libraries, "smart" homes) relying on a large number of diverse, interrelated data sources, but having no way to manage their *dataspaces* in a convenient, integrated, or principled fashion. This paper proposes dataspaces and their support systems as a new agenda for data management. This agenda encompasses much of the work going on in data management today, while posing additional research objectives.

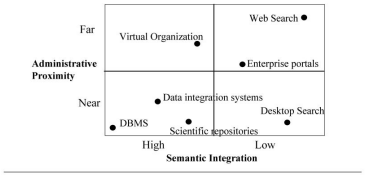


Figure 1. A space of data management solutions.

Source: Franklin, Halevy, and Maier (2005). From Databases to Dataspaces.

A brief history of data spaces

2010s

The concept shifted from academic data management to industrial and cross-organizational data sharing

A major milestone was the start of the **Industrial Data Space** initiative in Germany in 2014 by the **Fraunhofer Society**, which later evolved into the broader **International Data Spaces** movement.

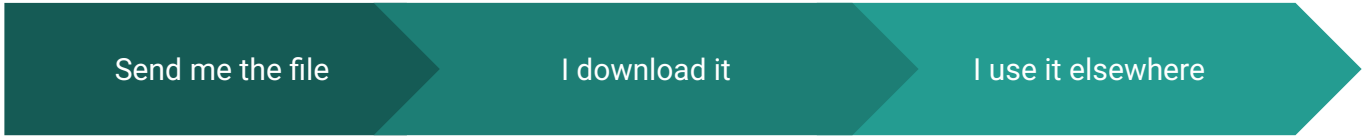
The **International Data Spaces Association (IDSA)** was created to define reference architectures and standards for sovereign data sharing.

Source: Wikipedia (2025). International Data Spaces Association.

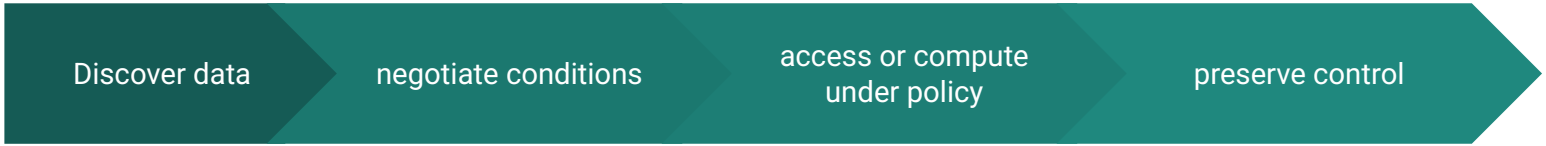
From data sharing to governed data sharing

Data spaces are not just repositories; they are governed ecosystems

- **Traditional data sharing often means:**



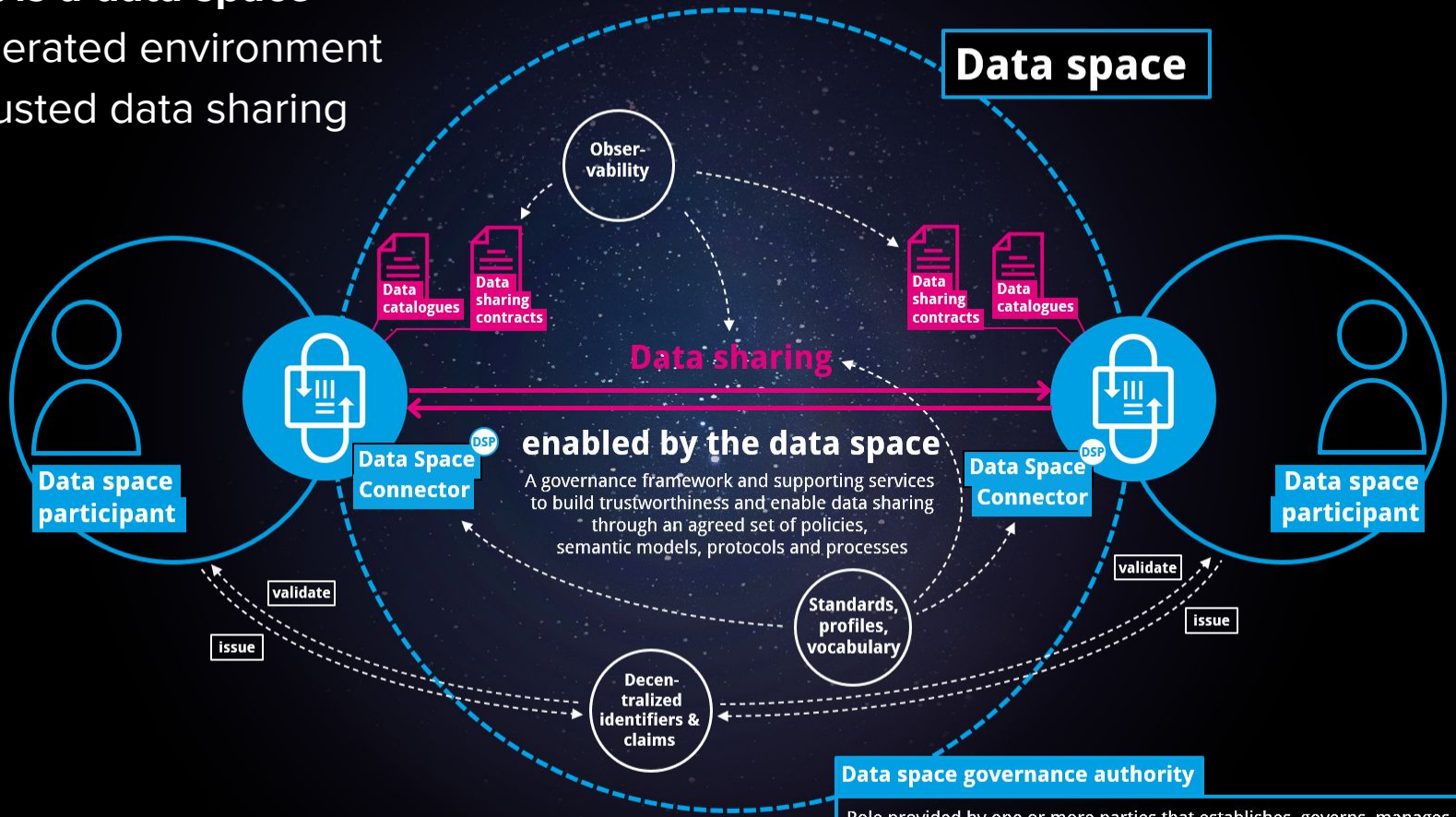
- **Data spaces aim for something different:**



Source: The author (2026).

What is a data space

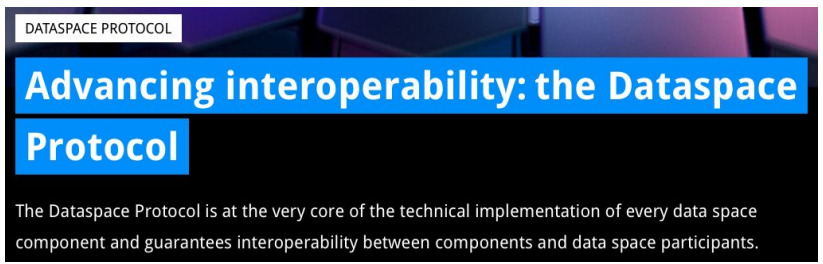
A federated environment for trusted data sharing



Source: International Data Spaces Association (2026). Data Spaces: Where the future of data happens.

The Main Players and Governing Rules

INTERNATIONAL DATA SPACES ASSOCIATION



European Commission

European Strategy for Data

A cross-sectoral legislative framework

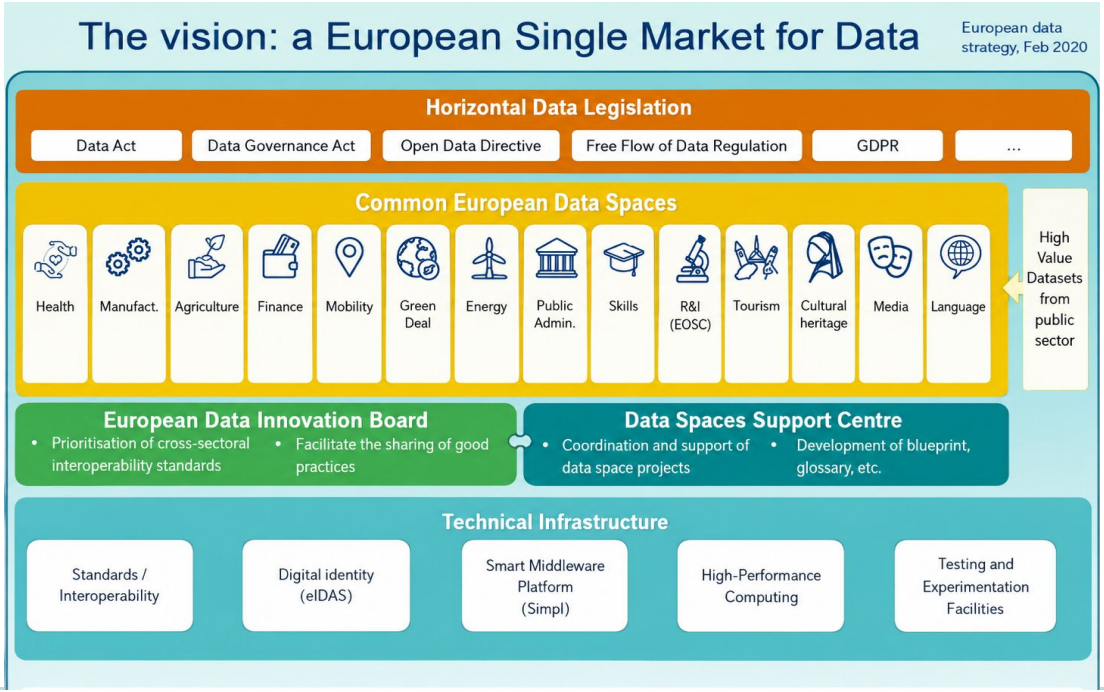
- **Data Governance Act**
- **Data Act**
- **Open Data Directive**
- **High-Value Datasets Regulation**
- **Interoperable Europe Act**
- **AI Act** (indirectly)

Sources: International Data Spaces Association (2026). Dataspace Protocol.
European Commission (2025). A European strategy for data.

European data spaces

Data spaces are a major part of the European data strategy

The European Union promotes Common European Data Spaces to enable trusted data sharing in strategic sectors.

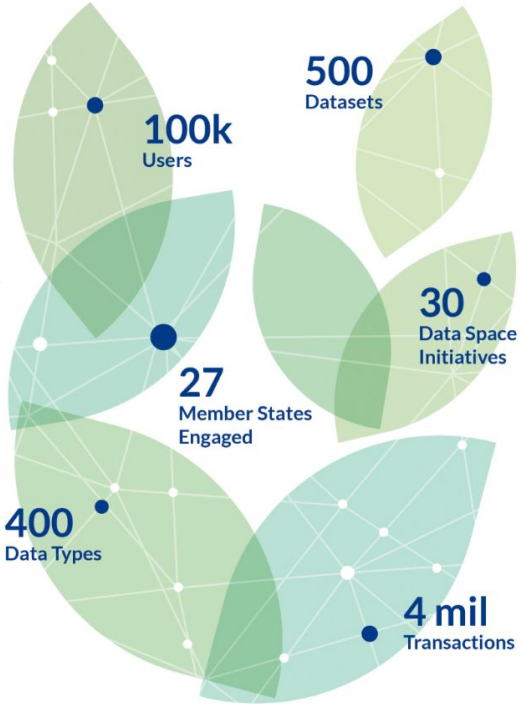


Source: Otto, Boris (2026). Smart Data Ecosystems.

The Common European Agricultural Data Space (CEADS)

What is it?

A federation of DSIs

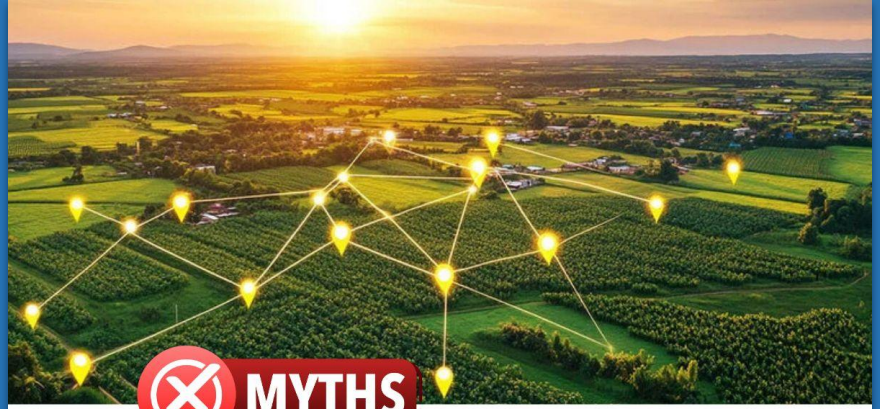


36 Partners

from 15 EU countries

A diverse consortium from 15 EU countries driving agricultural data innovation.

Source: CEADS Project (2026).



MYTHS

CEADS means sharing all farm data openly

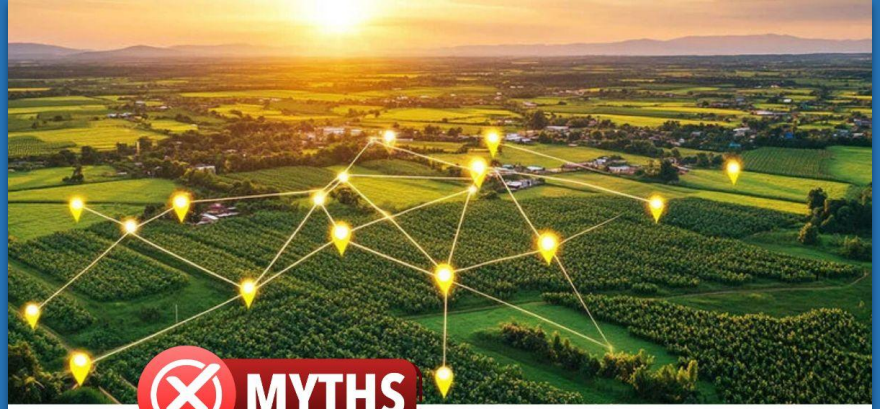
REALITY

CEADS is built on **data sovereignty.** Data owners (farmers, companies, organisations) decide if, how, and with whom their data is shared. Nothing is shared by default.

Source: CEADS Project (2026).

Co-funded by the European Union

CEADS
COMMON EUROPEAN AGRICULTURAL DATA SPACE



MYTHS

CEADS will centralise agricultural data in one big database

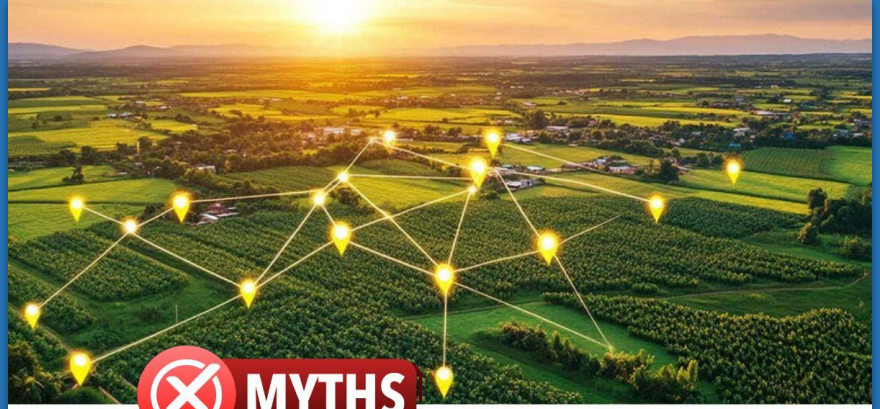
REALITY

CEADS is a **federated data space,** not a central database. Data stays where it is and is accessed securely via interoperable services and standards.

Co-funded by the European Union

CEADS

Source: CEADS Project (2026).



MYTHS

CEADS replaces existing platforms and systems

REALITY

CEADS connects and interoperates with existing platforms, data sources and services. It builds on what already exists instead of replacing it.

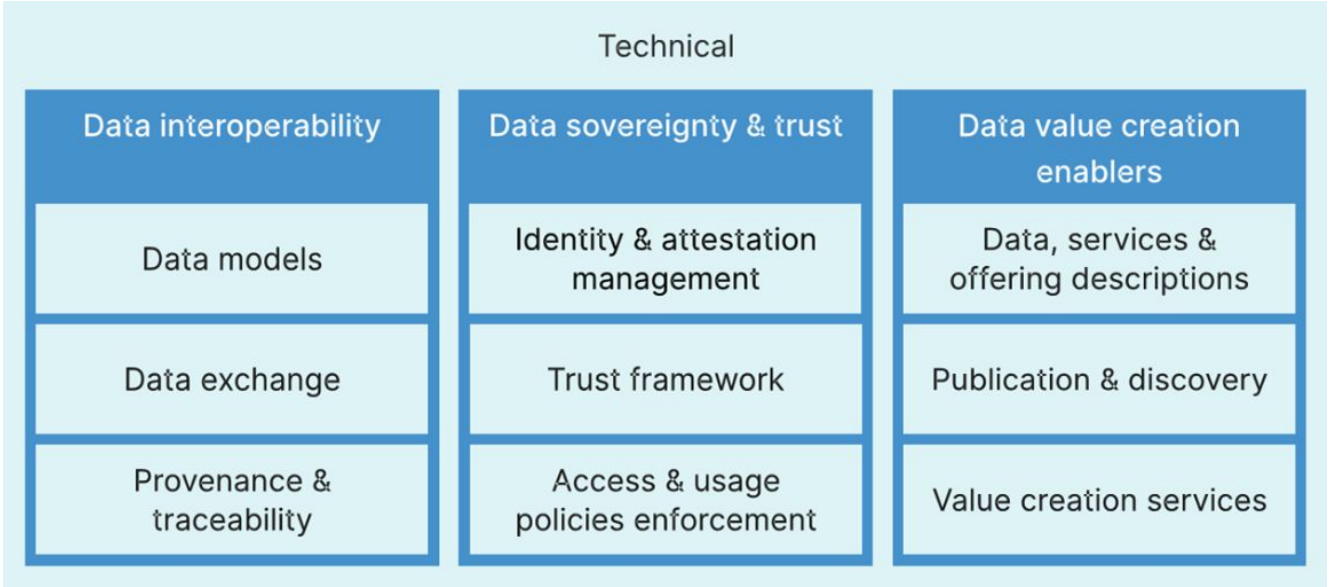
Co-funded by the European Union

CEADS

Source: CEADS Project (2026).

The Three Technical Pillars

(For all Data Spaces, CEADS included)



Source: Data Spaces Support Centre (2025). Technical Building Blocks.

Semantic Interoperability in the Data Spaces

Semantic interoperability is about preserving meaning across systems

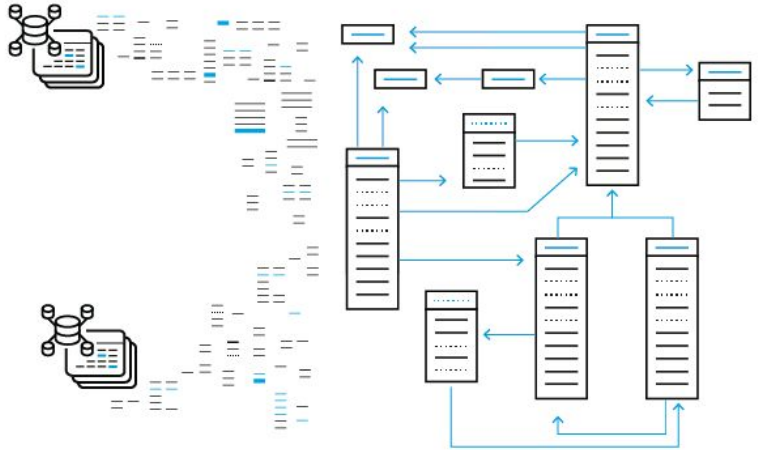
SEMIC supports semantic interoperability in Europe by promoting common data models, vocabularies, and reusable semantic assets

European metadata standards

- SEMIC is especially relevant because many European metadata standards and application profiles are connected to this ecosystem, including DCAT-AP and related profiles.

The data catalog vocabulary – application profile **DCAT-AP** is the application profile of the data catalogue vocabulary **(DCAT) ontology** created to increase interoperability and provide guidance to EU Member States that publish datasets on national (open) data portals.

DCAT-AP is a specification based on the World Wide Web Consortium's (W3C) DCAT for **describing the metadata** of public sector datasets in Europe.

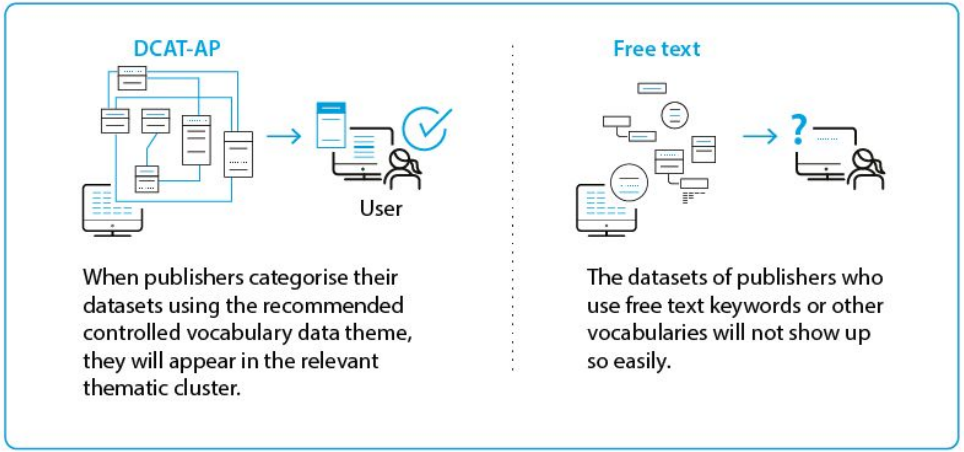


Source: European data (2022). Linking data: Data Catalogue Vocabulary Application Profile.

Semantic Interoperability in the Data Spaces

What are the specific features of DCAT-AP?

- DCAT-AP indicates **classes and properties that are mandatory**
- When describing properties, DCAT-AP r**equires the use of several EU controlled vocabularies**
- **It defines ranges and domains**. For example, a publisher of a dataset (domain) must be defined as belonging to the class of an agent (range)



Source: European data (2022). Linking data: Data Catalogue Vocabulary Application Profile.

Semantic Interoperability in CEADS

What role are we playing?

Organized in three scenarios

Scenario 1

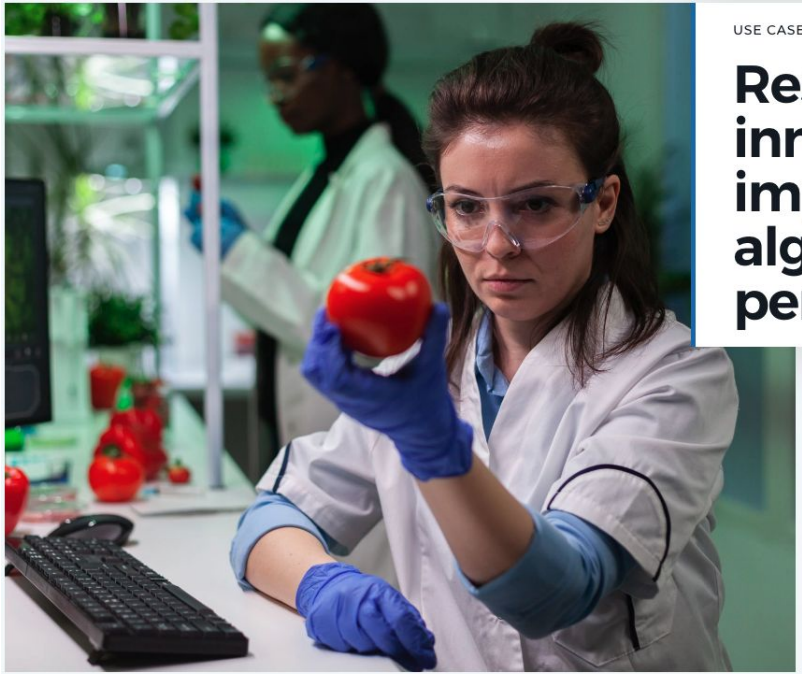
- Semantic Annotation at the Metadata Layer

Scenario 2

- AI-ready image datasets with semantic annotations

Scenario 3

- Harmonisation and semantic annotation at the data layer



USE CASE #2

Research and innovation data for improved AI algorithms performance

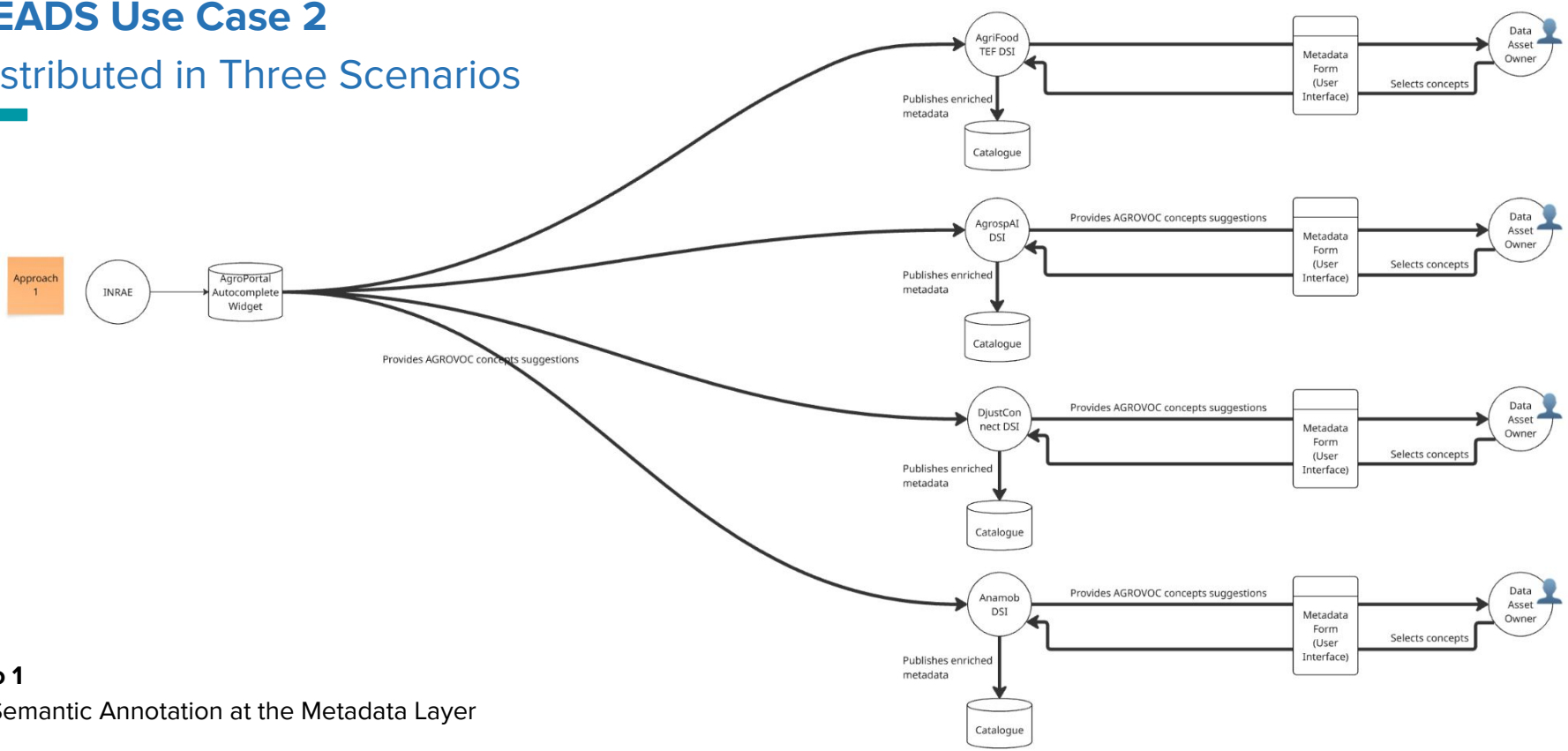
Countries: FR, IT, ES, BE, PL
Lead Partner: INRAE
Other Partners: IESE, FBK, UDL, EV ILVO, PSNC, ANAMOB, JR

This Use Case connects CEADS with the **AgrifoodTEF Data Space**, the upcoming **Agriculture of Data partnership**, and the **EOSC**. It ensures that datasets from experiments, robotics, and AI testing are accessible for research and business innovation.

Source: CEADS Project (2026). Use Case 2.

CEADS Use Case 2

Distributed in Three Scenarios



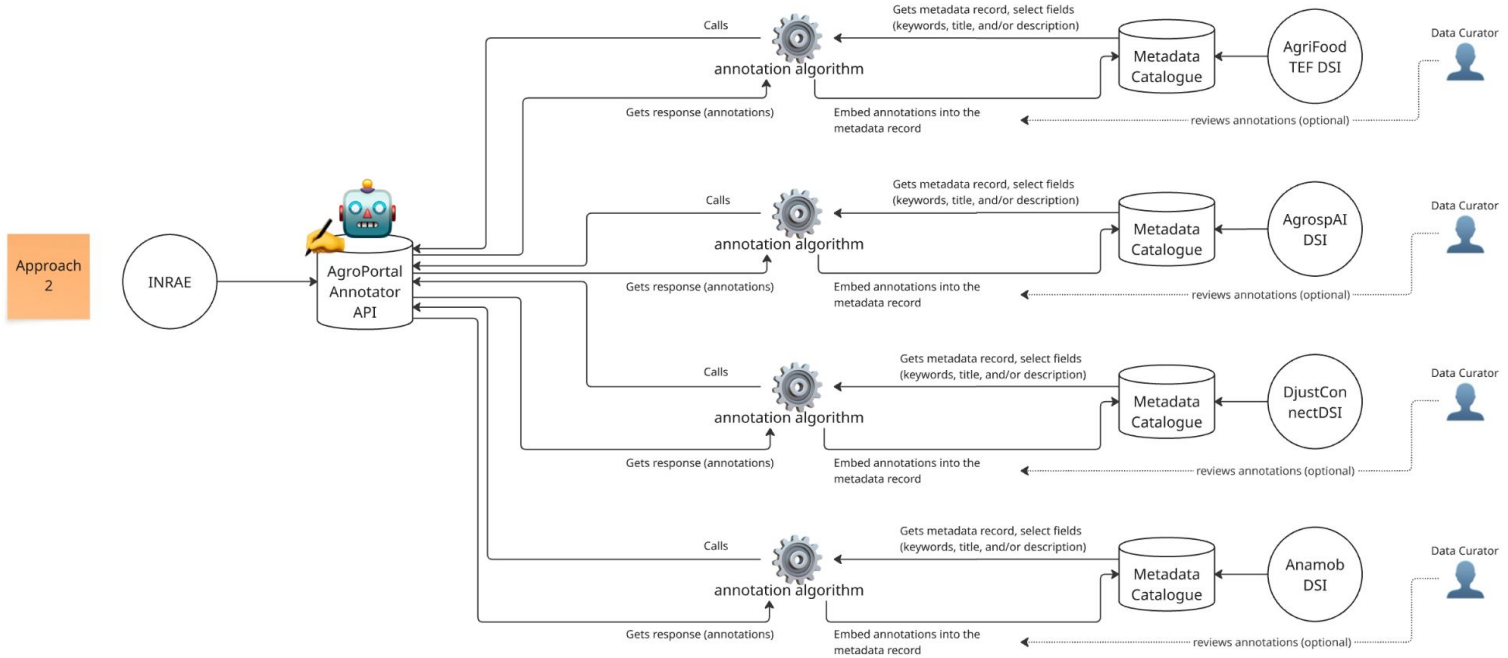
Scenario 1

- Semantic Annotation at the Metadata Layer

Source: The author (2026).

CEADS Use Case 2

Distributed in Three Scenarios



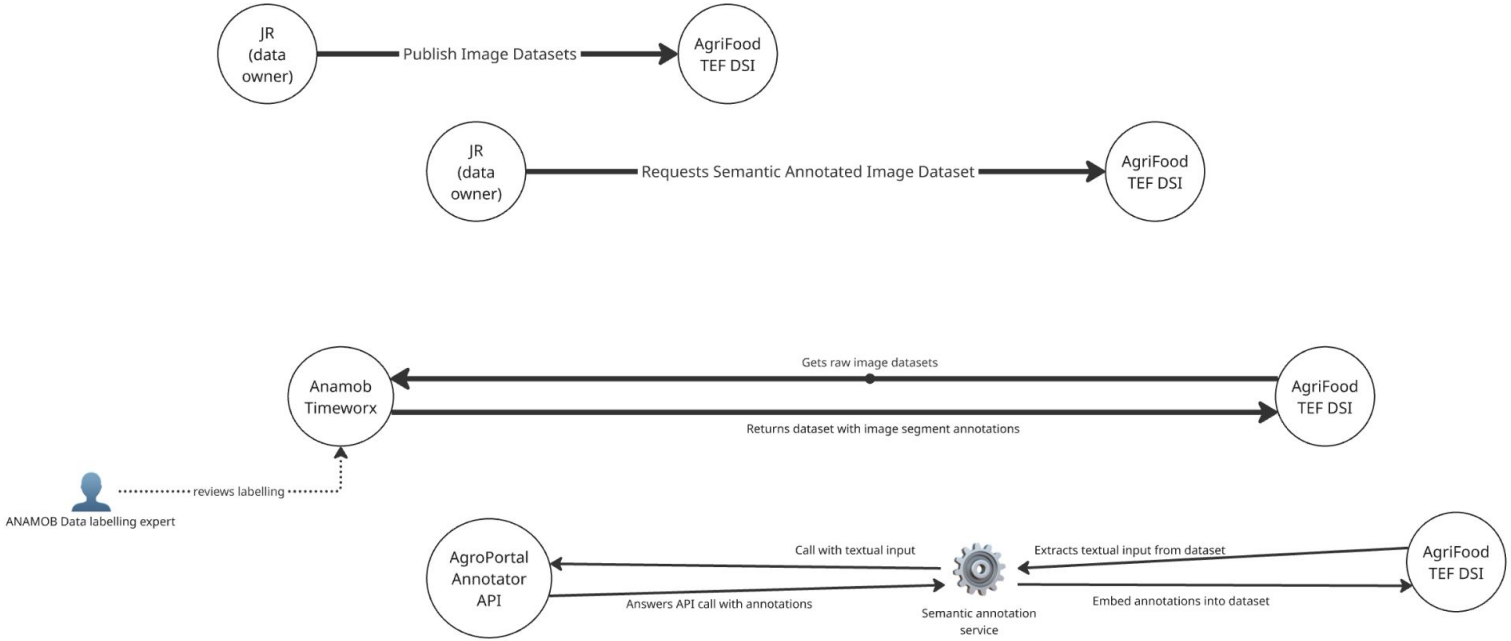
Scenario 1

- Semantic Annotation at the Metadata Layer

Source: The author (2026).

CEADS Use Case 2

Distributed in Three Scenarios



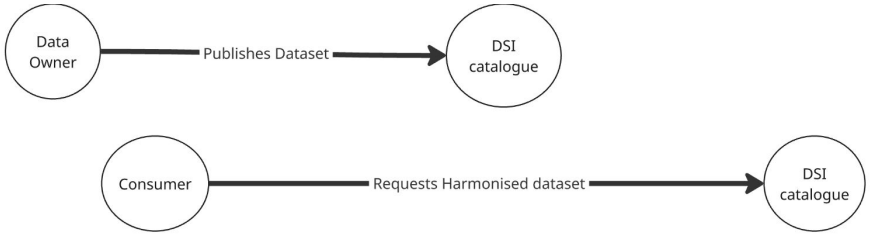
Scenario 2

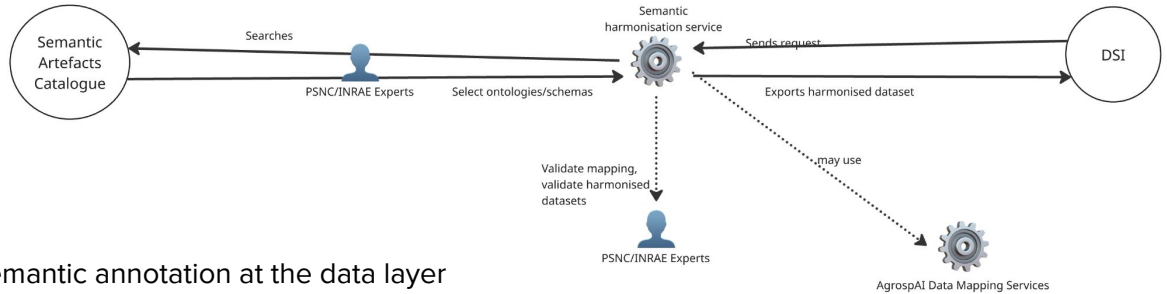
- AI-ready image datasets with semantic annotations

Source: The author (2026).

CEADS Use Case 2

Distributed in Three Scenarios





Scenario 3

- Harmonisation and semantic annotation at the data layer

Source: The author (2026).

CEADS Use Case 2

Where we currently are

Scenario 1: Three demonstrations with AgrospAI

1. AgroAnnotator
2. Connector from AgrospAI to AgroPortal
3. GeoDCAT-AP Wrapper

Source: AgrospAI. Universitat de Lleida (2025).

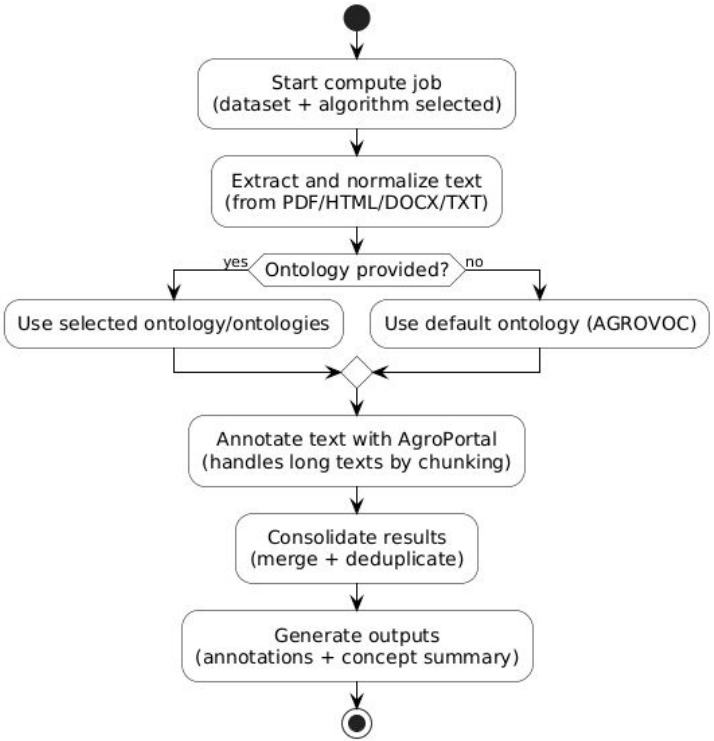
CEADS Use Case 2

Where we currently are

Scenario 1: Three demonstrations with AgrospAI

AgroAnnotator

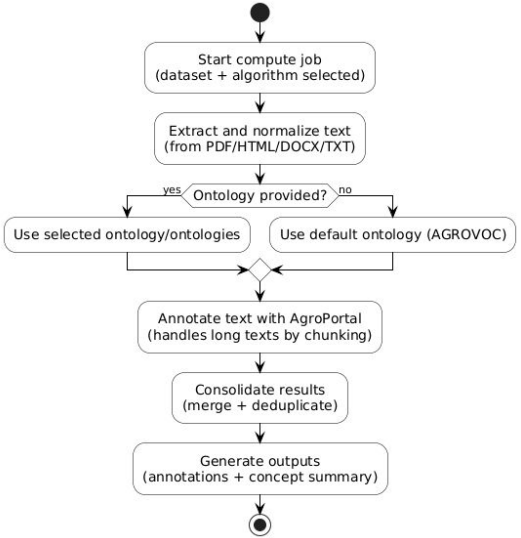
- AgroAnnotator is a Compute-to-Data algorithm written in Python 3.11 that performs ontology-based annotation of agricultural documents/datasets based on the AgroPortal Annotator service inside the secure AgrospAI execution environment.



Source: Soares, Jonquet, and Garcia (2026). Bringing Semantics into the Common European Agricultural Data Space (CEADS)

CEADS Use Case 2

Where we currently are



Job finished

Annotator Test Dataset (PDF) ↗
GXAT | did:op:1432270ffdfba2c5d316c99f5c98e12571dc0fb7d31a775

AgroPortal Annotator ↗
GXAT | did:op:668638a5d407bdd00188562ae188e544fc8937171e0b

Results

- RESULTS (chunk_0001.json) - 201.8 KB
- RESULTS (chunk_0002.json) - 180.1 KB
- RESULTS (chunk_0003.json) - 31.9 KB
- RESULTS (combined.json) - 461.7 KB
- RESULTS (concepts_summary.csv) - 12.7 KB
- RESULTS (concepts_summary.json) - 29.4 KB
- RESULTS (merged_annotations.json) - 440.8 KB
- RESULTS (run_metadata.json) - 581 Bytes
- ALGORITHM LOGS - 240 Bytes
- CONFIGURATION LOGS - 2.4 KB
- PUBLISH LOGS - 363 Bytes

Results are stored for 30 days.

CREATED
9 days ago

FINISHED
9 days ago

JOB ID
49eb7abc0e514a0ca72deac35539bb19

Source: Soares, Jonquet, and Garcia (2026). Bringing Semantics into the Common European Agricultural Data Space (CEADS)

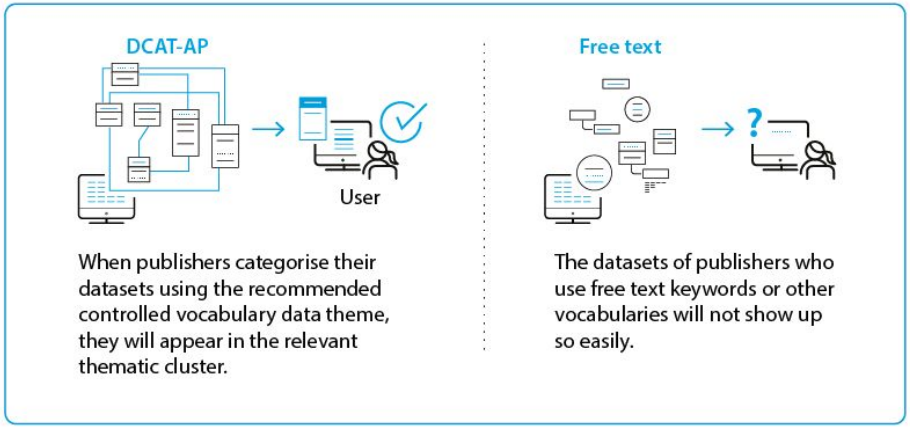
CEADS Use Case 2

Where we currently are

Scenario 1: Three demonstrations with AgrospAI

Connector from AgrospAI to AgroPortal

- An integration between AgrospAI and AgroPortal that enables tagging with controlled terms during asset publication



Source: Soares, Jonquet, and Garcia (2026). Bringing Semantics into the Common European Agricultural Data Space (CEADS)

CEADS Use Case 2

Where we currently are

Scenario 1: Three demonstrations with AgrospAI

Connector from AgrospAI to AgroPortal

(a) Ontology term selector fields in the publish form.

(b) Searching "wheat" without ontology filtering shows matches across multiple ontologies.

(c) Searching "wheat" with AGROVOC selected restricts suggestions to that ontology.

Source: Soares, Jonquet, and Garcia (2026). Bringing Semantics into the Common European Agricultural Data Space (CEADS)

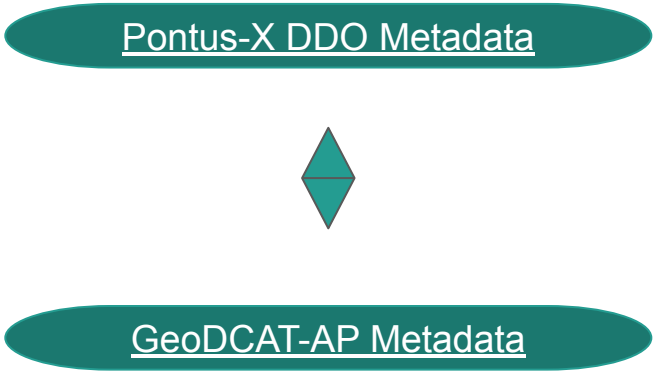
CEADS Use Case 2

Where we currently are

Scenario 1: Three demonstrations with AgrospAI

GeoDCAT-AP Wrapper

- A catalogue-level service that improves interoperability by exposing Pontus-X catalogue records as GeoDCAT-compliant metadata.



Source: Soares, Jonquet, and Garcia (2026). Bringing Semantics into the Common European Agricultural Data Space (CEADS)

CEADS Use Case 2

Discussion: what we learned

Lessons learned

- Data-space infrastructure alone is not enough.
- Metadata standards need practical mappings.
- Controlled vocabularies must be integrated into user workflows.
- Semantic services should respect data-space trust constraints.
- Compute-to-Data is promising, but external annotation APIs raise sovereignty questions.
- SHACL validation is essential for catalogue-level quality assurance.
- AgroPortal can support CEADS as a reusable semantic layer.

Source: Soares, Jonquet, and Garcia (2026). Bringing Semantics into the Common European Agricultural Data Space (CEADS)

CEADS Use Case 2

Future directions

What to do next

- integrate more ontologies beyond AGROVOC,
- support ontology selection by data providers,
- improve multilingual concept labeling,
- annotate structured datasets, not only documents,
- deploy annotation closer to the data-space environment,
- align with HVD DCAT-AP and other European profiles,
- generalize the connector to other Pontus-X installations

Source: Soares, Jonquet, and Garcia (2026). Bringing Semantics into the Common European Agricultural Data Space (CEADS)

Final message

Data spaces need semantics to become truly interoperable

Data spaces are not only about connecting infrastructures

They also require:

- shared meaning,
- semantic artefacts,
- controlled vocabularies,
- metadata standards,
- validation mechanisms,
- governance-aware semantic services.

AgroPortal can help CEADS move from:

data sharing

to:

meaningful, trusted, AI-ready data reuse

Source: Soares, Jonquet, and Garcia (2026). Bringing Semantics into the Common European Agricultural Data Space (CEADS)

Co-funded by the European Union
GA no: 101195295

CEADS
COMMON EUROPEAN AGRICULTURAL DATA SPACE

ceads@ilvo.vlaanderen.be
ceads.eu

Thank you for your attention!

Office on the web Frame