OneHealth DataSpace

CESGA
GALICIA SUPERCOMPUTING CENTER

# OneHealth DataSpace

## From sovereign data sharing to Big Data, HPC and AI across the full data lifecycle

PEDRO FERREIRO REGA · JAVIER CACHEIRO LÓPEZ

ON BEHALF OF ALL THE ONEHEALTH DATASPACE TEAM
CESGA

Financiado por la Unión Europea
NextGenerationEU

GOBIERNO DE ESPAÑA
MINISTERIO PARA LA TRANSFORMACIÓN DIGITAL Y DE LA FUNCIÓN PÚBLICA
SECRETARÍA DE ESTADO DE DIGITALIZACIÓN E INTELIGENCIA ARTIFICIAL

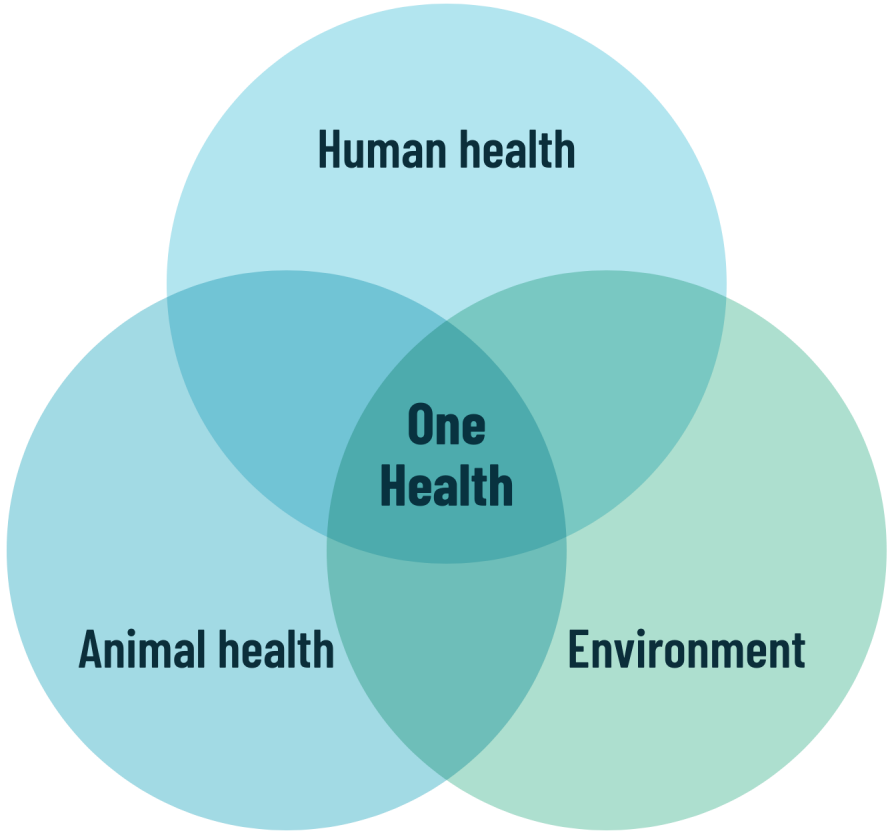
Plan de Recuperación, Transformación y Resiliencia

Xacobeo 2027

# What we will cover

1. One Health
2. Data Spaces
3. OneHealth DataSpace: evolution to cover the full data life cycle
4. Use Cases
5. Demo

# One Health




WHAT IS ONE HEALTH?

An **integrated** approach that recognises that the health of people, animals and the environment is closely **linked** and **interdependent**

THE OPPORTUNITY

# Creating value from One Health data

## ONE HEALTH DATA IS ALREADY THERE

- Data of many kinds with enormous potential
- Organisations with specialised knowledge

↓

- **The One Health ecosystem has a lot of potential to create value from data**

## THE CAPABILITIES ARE HERE

### HPC
Computing power for complex analyses and simulations

### Big Data
Large-scale data analysis with intelligent models

### Cloud
Scalable, accessible infrastructure to expose results

# The barriers

Why is not happening?
Because many organisations…

- **Need to keep sovereignty over their data**
- **Lack advanced infrastructure**
- **Do not have all the specialised profiles**
- **Have data, but it is not always usable**

# Data Spaces: Collaborating without losing control

## The data space concept

A structured framework for exchanging, accessing, managing and using data across multiple entities in a secure and interoperable way

- **Data exchange**
- **Data sovereignty**
- **Interoperability**
- **Governance and trust**

## Analogy: the food market



THE KEY COMPONENT

# The connector

## Adds sovereignty

- Attaches usage policies to each asset
- Negotiates contracts with other participants
- Enforces them before any data leaves




One per participant

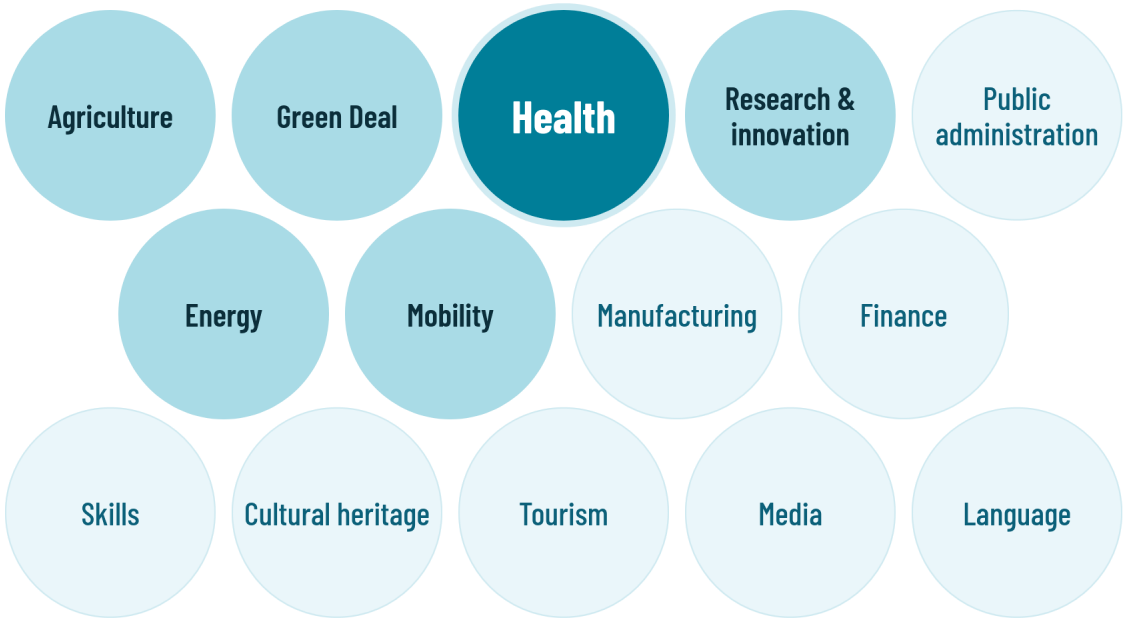
## Enables federation

- Publishes to and searches the shared catalogue
- Verifies the identity of other participants
- Exchanges data directly, with no central store

# Common European Data Spaces

The European strategy for data defines sector-specific data spaces that share common rules and principles, so data can flow across organisations and borders

SECTOR SPECIFIC




REGULATORY FRAMEWORK

## European Health Data Space

`Regulation (EU) 2025/327`

**Primary use:** people access and control their health data across the EU

**Secondary use:** reuse for research, innovation and policy in secure environments

| | |
|---|---|
| **Data Governance Act** | Trusted data intermediation |
| **Data Act** | Fair access to and use of data |
| **AI Act** | Risk-based rules for trustworthy AI |

# Evolving the Data Space Concept: OneHealth DataSpace

ONEHEALTH DATASPACE ADDS

## CESGA's advanced capabilities

to build real solutions from shared data

CESGA
GALICIA SUPERCOMPUTING CENTER

BUILT ON TOP OF A CONVENTIONAL DATA SPACE

- ✓ Data exchange
- ✓ Interoperability
- ✓ Governance
- ✓ Data sovereignty

ADVANCED CAPABILITIES BUILT INTO OHDS

- Data analytics
- Advanced computing
- Local AI and LLMs
- Big Data
- Scalable storage
- Service delivery

## More than data

# One platform, the full data lifecycle

**Share**
Share data securely with other participants

**Store**
Store and protect data assets

**Analyze**
Turn data into knowledge

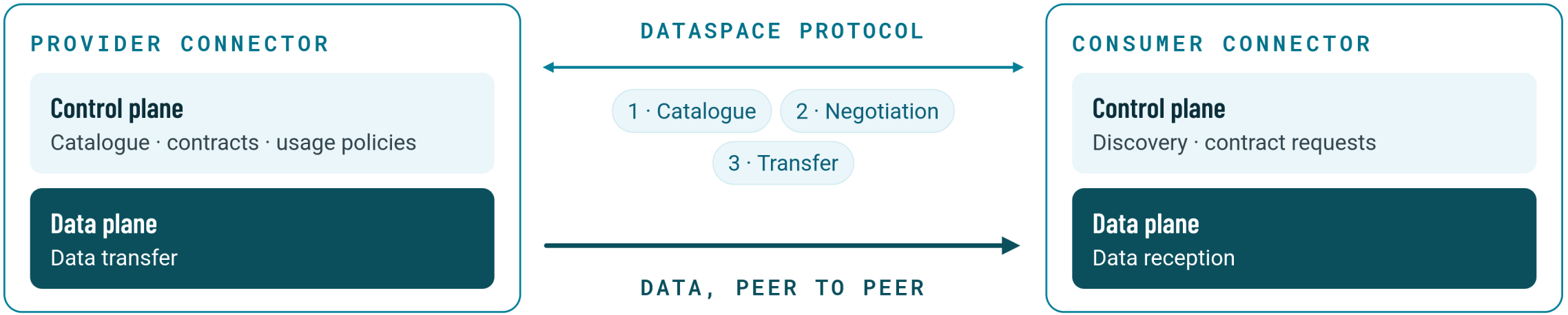
**Compute**
Use advanced computing capabilities

**Deliver**
Deliver results as services and applications



# Data Space Middleware: Eclipse Dataspace Components

Every participant runs a connector. EDC is the open-source framework from the Eclipse Foundation



# How to connect to OHDS

## Local connector

**Full data sovereignty**

- Data never leaves your infrastructure
- Full technical and operational autonomy

**Simplified deployment**
Automated installation, minimal effort

## Connector as a service

**No infrastructure needed**

- CESGA deploys and maintains it for you
- Onboarding to the data space in minutes

**Ideal to get started**
For entities without own infrastructure

# Governance




- Common rules for access, use and quality
- Clear roles: data and service providers and consumers
- Regulatory compliance

MEMBERSHIP PROCESS

1

**Learn about OHDS**

Download the Participant Guide: technical and legal requirements and benefits

2

**Apply to join**

Expression of interest form, reviewed in the evaluation cycle

3

**Technical integration**

Connect your systems and prepare your data to meet interoperability and security requirements

4

**Get value**

Exchange data, join use cases and collaborate in the ecosystem

# Membership Agreement

Before joining, every participant signs a membership agreement and accepts the data space's common governance rules

## What each participant commits to

**Rules of use**
How data can be accessed and used

**Roles**
Duties of providers and consumers

**Compliance**
Data protection and sector regulation

**Accountability**
What happens if the rules are broken

HOW IT BUILDS TRUST

**Technical trust** — Credentials and connectors prove who each participant is

+

**Legal trust** — The agreement binds every member to the same commitments

You can share data with partners you have never worked with before

# Sharing data: assets

The provider decides what to share, with whom and under which conditions. Nothing is visible in the catalogue unless the provider allows it

1 · ASSET

## Register what you share

A dataset, file, API or service, with its description and location

Private by default

2 · POLICIES

## Set the rules

**Access:** who can see the offer

**Usage:** conditions to use the data

3 · CONTRACT DEFINITION

## Link asset and policies

Only now the asset appears in the catalogue, and only for the participants you allow

Listed in the catalogue

4 · CONTRACT AGREEMENT

## Negotiate and transfer

The consumer accepts the offer, both connectors sign the agreement and the data can be transferred

# The catalogue

Each connector publishes what it offers using DCAT, the W3C vocabulary for describing datasets, so any participant can read any catalogue

WHAT A CONNECTOR RETURNS

`dcat:Catalog` the participant's offer

- `dcat:DataService` the connector endpoint
- `dcat:Dataset` one per asset: title, description, keywords
  - `odrl:hasPolicy` the terms of use offered

Serialised as JSON-LD · exchanged over the Dataspace Protocol

### DCAT describes the data

Standard metadata that any connector can interpret

### ODRL describes the rules

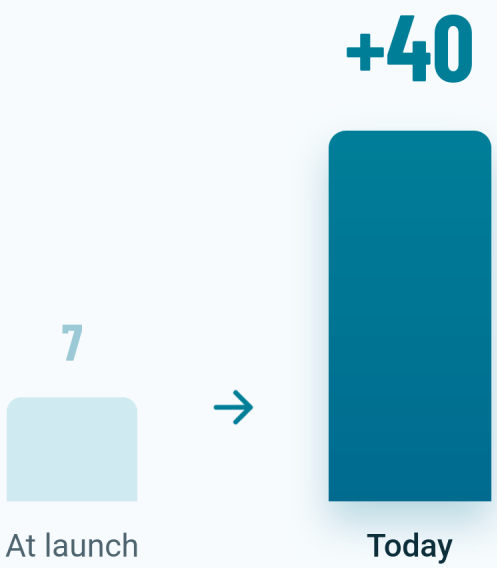
Each offer carries the policy that will be negotiated

### Federated catalogue

Periodically collects every participant's catalogue, so the whole data space can be searched in one place

# Current status

## USE CASE GROWTH




**6 times more** use cases since launch

## PARTICIPANTS

**36** entities have joined OHDS



36 members ≈30 in the process of joining



| | |
|---|---|
| Companies | 20 |
| Research/Univ. | 12 |
| Public admin. | 2 |
| Association/NGO | 2 |

**International presence**
Active collaborations with institutions abroad

**All OneHealth sectors represented**
Full coverage, with human, animal and environmental health use cases

# Some OHDS use cases

## HUMAN HEALTH

MD.USE Innovative Solutions S.L.

### CD-Ligand ML: Molecular modelling and AI for bioactive ligands

Drug discovery: Interaction with functional encapsulation systems

GHX

### Large-scale virtual screening on EGFR

Virtual screening: GPU-accelerated docking and molecular dynamics

MÁS MÉTODO

### Territorial health profiles with AI and climate

Comparable profiles of basic health areas for health planning

de la Fuente Lab

### Evo 2: Fine-tuning with viral genomes

Using LoRA on models of different sizes

## ENVIRON-MENTAL HEALTH

INSTITUTO ESPAÑOL DE OCEANOGRAFÍA

### Portuguese man o' war on Galician beaches

Daily drift simulation with currents, wind and HF radar

METEOGALICIA

### WRF-Fire: Wildfire modelling

Coupled atmosphere-fire simulation from the 1 km WRF

## ANIMAL HEALTH

consuvet CONSULTING VETERINARIO

### CcVet: Digitising veterinary clinical cases for data spaces

Intersecting with human and environmental health

# Use cases (SMEs)

imatia

### DataBench: a unified view of distributed data

**IMATIA INNOVATION**

Consolidates clinical records fragmented across multiple systems

COGNITUS

### AID-G: an AI avatar against unwanted loneliness

**COGNITUS BIOTECH**

Anonymised metrics from conversational companionship for older people

### Digitalising day centre activity

**OTTO SOFTWARE**

Real-time care practice data from healthcare day centres

INSATI

### VAIRA: physical-cognitive stimulation

**INSATI INNOVATION**

Physical-cognitive stimulation activity data for R&D

Nexthealth

### 1MillionHCE: synthetic health records

**NEXT & HEALTH CONSULTING**

Synthetic, longitudinal electronic health records under the IPS standard

Mercantia

### MicroBioToolDB: human gut microbiome

**MERCANTIA DESARROLLOS ALIMENTARIOS**

Curated database for personalised nutrition research

# Use cases (SMEs)

TRUEWORLD

### Environmental and agronomic data from IoT sensors

TRUEWORLD

Weather stations and farm sensors applied to environmental health

### MatLabID: open catalogue of lab materials

ICIGA DATALAB

Reusable catalogue of lab equipment, consumables and reagents

MÁS MÉTODO

### Territorial health profiles with AI and climate

MÁS MÉTODO CONSULTORÍA

Comparable profiles of basic health areas for health planning

Agrolinera

### IoT whey management in Cabrales cheese dairies

AGROLINERA ASTUR

Traceability and circular economy in artisan mountain cheese dairies

TOPHEALTHTECH AI SOLUTIONS

### Clinical data of diabetic patients

TOP HEALTH TECH

Real clinical records to train and validate AI models for diabetes

### And many more...

Explore the full OHDS use case catalogue

dataspace.cesga.es
/es/catalog

# Conclusions

1 One Health data is available and has great potential to generate value

2 We have extended the data space concept to cover the full data lifecycle

3 Our main focus has been to help the use cases succeed

## How to join

dataspace.cesga.es/es/how

onehealth@cesga.es

What really matters is not the infrastructure we have built, but **the valuable results it makes possible**

Financiado por la Unión Europea
NextGenerationEU

GOBIERNO DE ESPAÑA
MINISTERIO PARA LA TRANSFORMACIÓN DIGITAL Y DE LA FUNCIÓN PÚBLICA
SECRETARÍA DE ESTADO DE DIGITALIZACIÓN E INTELIGENCIA ARTIFICIAL

Plan de Recuperación, Transformación y Resiliencia

Xacobeo 2027

LIVE DEMO

# Demo: OneHealth DataSpace in Action

dataspace.cesga.es/en ↗

