

Scaling Cross-Project Collaboration across AI4EOSC, SIESTA, and TITAN

Álvaro López García and Judith Sáinz-Pardo Díaz

on behalf of the EOSC SIESTA and AI4EOSC consortiums

aloga@ifca.es, sainzpardo@ifca.es

Advanced Computing and e-Science group
Institute of Physics of Cantabria (IFCA) CSIC-UC



**Funded by
the European Union**



AI4 |

- HORIZON-INFRA-2021-EOSC-01-04
- 5M€, grant based
- Finished: Sep 2022 – Aug 2025
- 9 partners (coord: CSIC)

Goal: Deliver an **EOSC AI platform with advanced features** for distributed, **federated learning**, metadata provenance, MLOps and provision of **AI/ML/DL services**.



AI platform, catalog of
models and advanced AI
tools

 | SIESTA

- HORIZON-INFRA-2023-EOSC-01-06
- 5M€, lump-sum scheme
- Ongoing: Jan 2024 – Dec 2026
- 12 partners (coord: CSIC)

Goal: Deliver **trusted cloud-based environments** for the **analysis of sensitive data**, and a set of services to ease the secure sharing through PETs and anonymization techniques.



Secure computing
platform, confidential
storage

 | TITAN

- HORIZON-INFRA-2023-EOSC-01-06
- 5M€, lump-sum scheme.
- Ongoing: Feb 2024 – Jan 2027
- 16 partners (coord: UMU)

Goal: Enriching the **EOSC Interoperability Framework** by developing a software platform solution for **confidential collaboration and privacy-preserving data processing**.

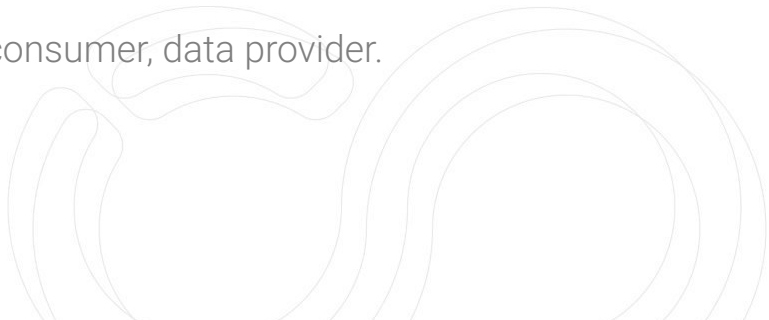


Decentralized data
exchange, governance
models

Use case

Goal: train an AI model for powdery mildew risk prediction in vineyards with private data distributed in different locations

- Powdery mildew is a fungal disease that causes significant yield and quality losses in viticulture.
 - Early risk prediction based on meteorological and phenological features
- Data is naturally distributed by Protected Designation of Origin (DPO).
 - Each data owner/client trains locally without sharing raw observations. A Federated learning architecture is adopted.
 - Data can be accessed via TITAN data space
 - Decentralized data exchange: allows to securely share, manage, and utilize data while retaining full ownership and control.
 - Roles: data owners, data producers, data consumer, data provider.
- Data and use case provider:



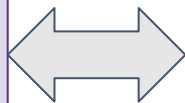
Problem statement

Challenge

Distributed data



Each data owner maintains data sovereignty, and accountability



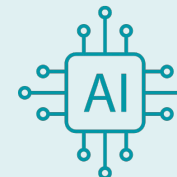
Privacy constraints



Each client keeps the data locally, without sharing or centralizing

Requirement

Development of an accurate AI model



The AI model needs to be trained with as much data as possible

Problem statement

Challenge

Distributed data



Each data owner maintains data sovereignty, and accountability



 |  TITAN

Privacy constraints



Each client keeps the data locally, without sharing or centralizing

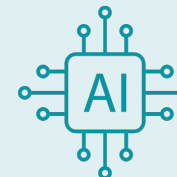


 |  SIESTA



Requirement

Development of an accurate AI model



The AI model needs to be trained with as much data as possible



AI4 | 

Federated learning:

*Bring computation to the data location
and train global AI models without data sharing*

The AI4 |  eosc platform provides users with a **federated learning server** that enables federated training with clients deployed on external cloud machines, locally, or within the same platform in different deployments.

The  eosc |  SIESTA platform provides users with a generic instance for deploying **federated learning clients** for performing the model training, exploiting the **secure storage**. These clients can connect to the server deployed in AI4EOSC

TITAN: data space

TITAN Data Space





TITAN



Data distributed
available in
TITAN data space

Assets Catalog

New Asset

Select a service, model, or dataset from our Catalog

Services

Models

Datasets

Explore and obtain datasets from different researchs and fields



Vinos de la Tierra

This dataset supports federated disease-risk prediction for vineyards in Vinos de la Tierra. It combines phytosanitary field...

Agriculture

Access



Somontano

This dataset supports federated disease-risk prediction for vineyards in Somontano. It combines phytosanitary field...

Agriculture

Access



Cariñena

This dataset supports federated disease-risk prediction for vineyards in Cariñena. It combines phytosanitary field...

Agriculture

Access



Calatayud

This dataset supports federated disease-risk prediction for vineyards in Calatayud. It combines phytosanitary field...

Agriculture

Access



Borja

This dataset supports federated disease-risk prediction for vineyards in Borja. It combines phytosanitary field...

Agriculture

Access

OR

AI4



Federated learning with Flower

Run federated learning trainings using the Flower framework.

Generic
server




SIESTA

Generic client

Client-fl

Client for the federated learning architecture. The client can connect to a server deployed on the AI4EOSC platform using Flower. This application enables distributed training with collaboration between multiple clients without sharing raw data.

Launch

Custom client
(TITAN)

Client-titan

Federated Learning clients for powdery mildew risk prediction in vineyards. Each client trains locally on data from a single Protected Designation of Origin (PDO) without sharing raw observations. Model weights are exchanged with a central server deployed in the AI4EOSC platform.

Launch

AI4 | 

Dashboard
Statistics
Catalog
Modules
Tools
LLMs
Tasks
Try me
Deployments
Batch training
Inference
Platform Services
Storage
Experiment tracking
AI Inference workflows
LLM Chat

Tools

Search

Additional filters
0

Sort:
By name
By most recent

6 result(s) found

Libraries
Tasks
Platform Categories
Data Types
Tags
Add filter

 **AI4life model loader**
Support for inference of the AI4LIFE model on the marketplace.

 **AI4OS Development Environment**
This is a Docker image for developing new modules

 **CVAT Image Annotation**
Image annotation tool that uses the Computer Vision Annotation Tool (CVAT)

 **Deploy your LLM**
Deploy your own LLM instance with VLLM and OpenWebUI.

 **Federated learning with Flower**
Run federated learning trainings using the Flower framework.

 **Federated learning with NVFlare**
NVFlare Dashboard and Federated learning server

The AI4EOSC dashboard is a service provided by CSIC, co-funded by [AI4EOSC](#)

v3.24.2
Status
Homepage
Docs
Terms of use
Privacy policy

Licensed under the Apache License 2.0. © 2026



SIESTA Dashboard

Project

30f55aef-88e6-4f2b-9322-8cb749eadbd5 personal project ▼

Privacy
PolicyAcceptable Use
Policy

Documentation

Project
Website

Logout



Reduce



Home



My account



Project settings



Service catalog



My Services



My Secrets



File Explorer



Data Explorer



Service catalog



Search



All

Anonymization
CatalogGeneral tools
CatalogEpidemiology
CatalogEnergy forecasting
CatalogEnergy forecasting Admin
CatalogText Anonymization
CatalogMedical Imaging
Catalog

Jupyter-python

The JupyterLab IDE with Python and a collection of standard data science packages.

Launch



Ubuntu22

Ubuntu 22.04 (Jammy Jellyfish) exposed via Wetty web-based terminal emulator.

Launch



Xfce-desktop

Lightweight XFCE desktop with VNC endpoint for Apache Guacamole

Launch



Guacamole

Apache Guacamole remote desktop gateway (guacamole + guacd + postgres)

Launch



Client-titan

Federated Learning clients for powdery mildew risk prediction in vineyards. Each client trains locally on data from a single Protected Designation of Origin (PDO) without sharing raw observations. Model weights are exchanged with a central server deployed in the AI4EOSC platform.

Learn more

Launch



Client-fl

Client for the federated learning architecture. The client can connect to a server deployed on the AI4EOSC platform using Flower. This application enables distributed training with collaboration between multiple clients without sharing raw data.

Launch





SIESTA Dashboard

Project

30f55aef-88e6-4f2b-9322-8cb749eadbd5 personal project ▾

Privacy
PolicyAcceptable Use
Policy

Documentation

Project
Website

Logout



Reduce



Home



My account



Project settings



Service catalog



My Services



My Secrets



File Explorer



Data Explorer



My Services



Access your running services



Services are supposed to be shut down as soon as you stop using them actively.



Refresh



New service



Delete all



Events

Running services

\$ helm list --namespace user-30f55aef-88e6-4f2b-9322-8cb749eadbd5



Client-calatayud ✎



Client-somontano ✎

Service

Status

Client-titan

Container starting ↻



Open

Service

Status

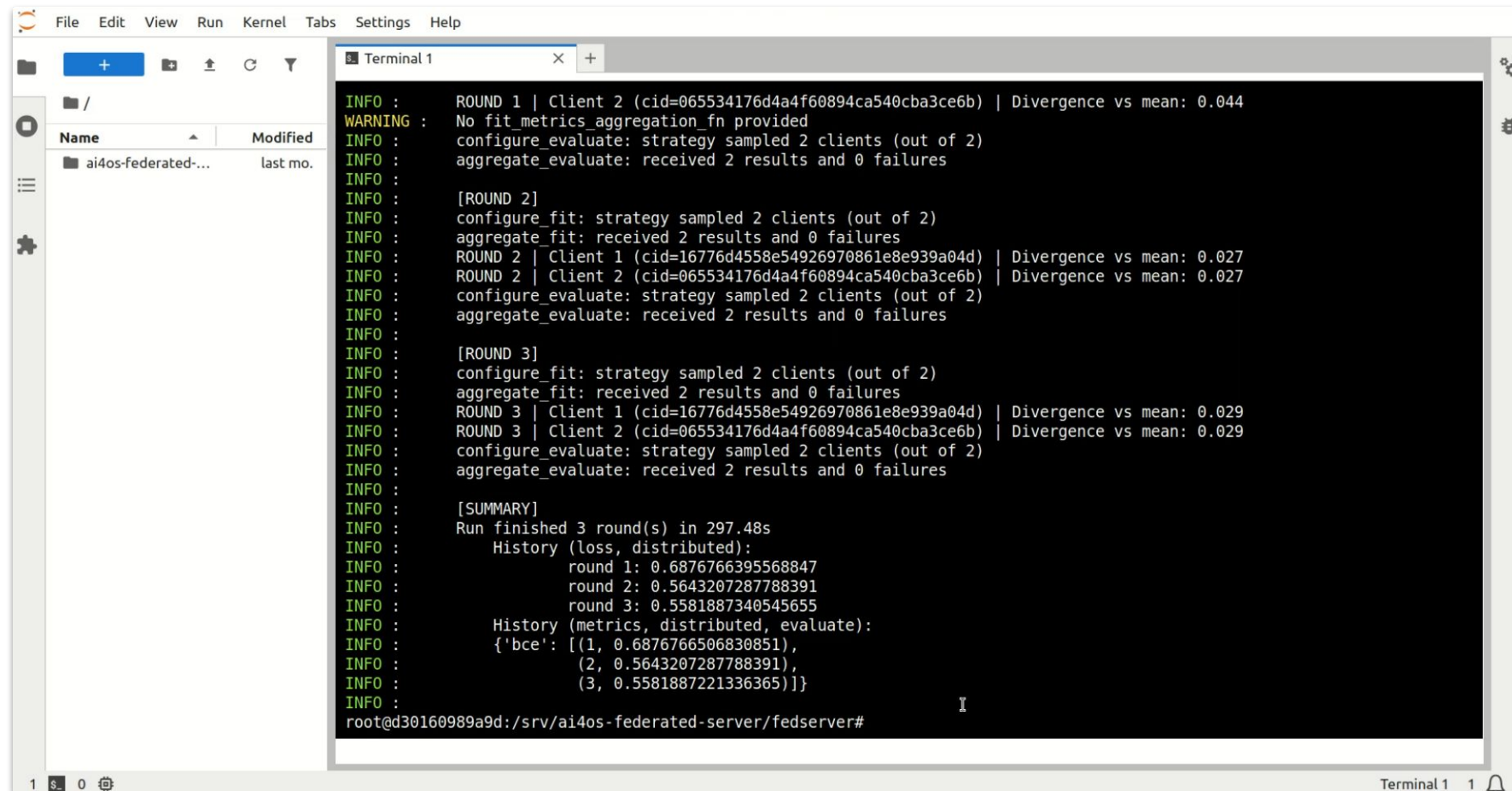
Client-titan

Container starting ↻



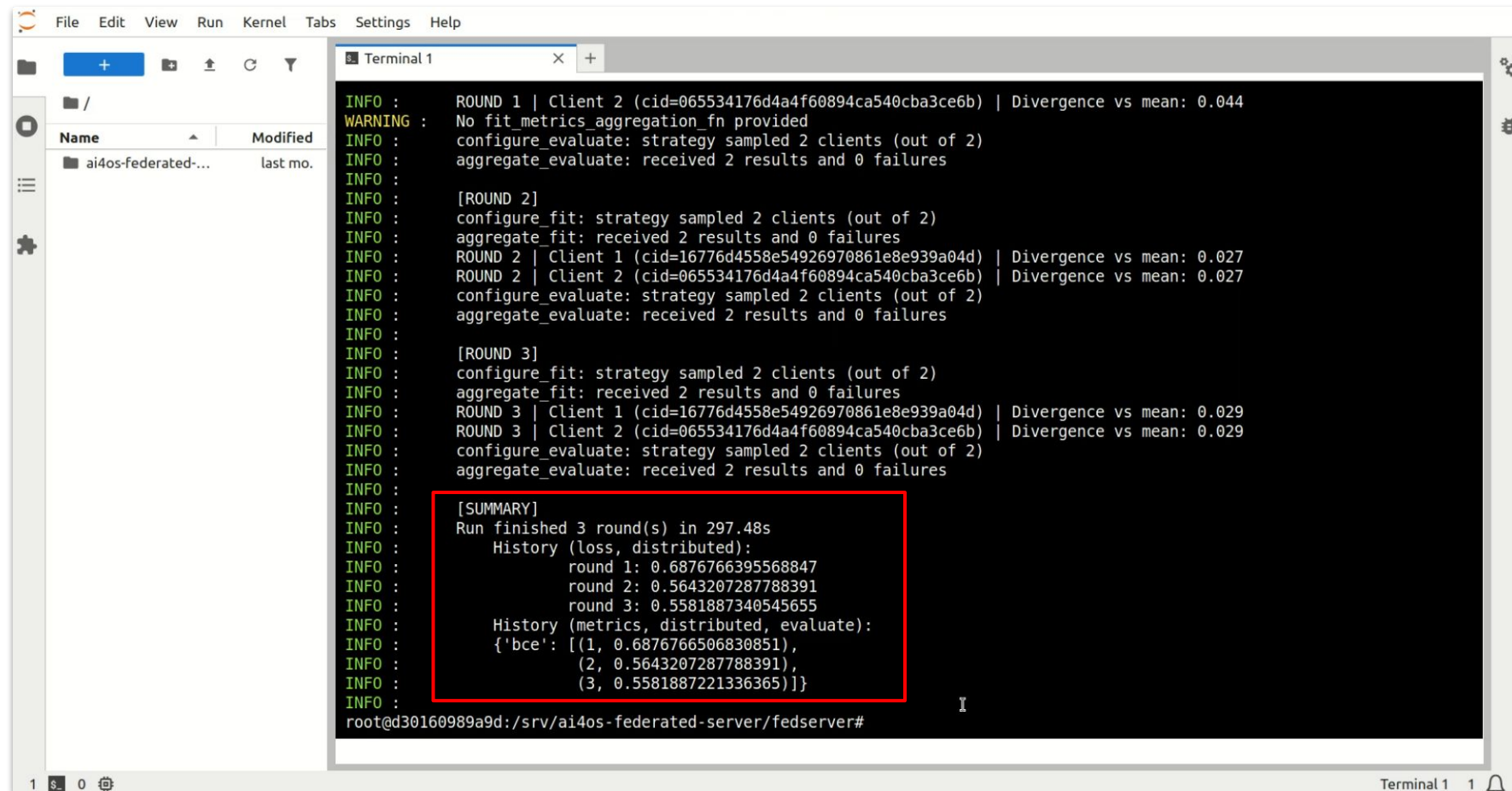
Open





```

INFO : ROUND 1 | Client 2 (cid=065534176d4a4f60894ca540cba3ce6b) | Divergence vs mean: 0.044
WARNING : No fit_metrics_aggregation_fn provided
INFO : configure_evaluate: strategy sampled 2 clients (out of 2)
INFO : aggregate_evaluate: received 2 results and 0 failures
INFO :
INFO : [ROUND 2]
INFO : configure_fit: strategy sampled 2 clients (out of 2)
INFO : aggregate_fit: received 2 results and 0 failures
INFO : ROUND 2 | Client 1 (cid=16776d4558e54926970861e8e939a04d) | Divergence vs mean: 0.027
INFO : ROUND 2 | Client 2 (cid=065534176d4a4f60894ca540cba3ce6b) | Divergence vs mean: 0.027
INFO : configure_evaluate: strategy sampled 2 clients (out of 2)
INFO : aggregate_evaluate: received 2 results and 0 failures
INFO :
INFO : [ROUND 3]
INFO : configure_fit: strategy sampled 2 clients (out of 2)
INFO : aggregate_fit: received 2 results and 0 failures
INFO : ROUND 3 | Client 1 (cid=16776d4558e54926970861e8e939a04d) | Divergence vs mean: 0.029
INFO : ROUND 3 | Client 2 (cid=065534176d4a4f60894ca540cba3ce6b) | Divergence vs mean: 0.029
INFO : configure_evaluate: strategy sampled 2 clients (out of 2)
INFO : aggregate_evaluate: received 2 results and 0 failures
INFO :
INFO : [SUMMARY]
INFO : Run finished 3 round(s) in 297.48s
INFO : History (loss, distributed):
INFO : round 1: 0.6876766395568847
INFO : round 2: 0.5643207287788391
INFO : round 3: 0.5581887340545655
INFO : History (metrics, distributed, evaluate):
INFO : {'bce': [(1, 0.6876766506830851),
INFO : (2, 0.5643207287788391),
INFO : (3, 0.5581887221336365)]}
INFO :
root@d30160989a9d:/srv/ai4os-federated-server/fedserver#
  
```



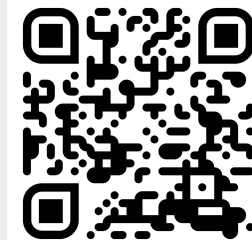
```

File Edit View Run Kernel Tabs Settings Help

/
Name Modified
ai4os-federated-... last mo.

Terminal 1

INFO : ROUND 1 | Client 2 (cid=065534176d4a4f60894ca540cba3ce6b) | Divergence vs mean: 0.044
WARNING : No fit_metrics_aggregation_fn provided
INFO : configure_evaluate: strategy sampled 2 clients (out of 2)
INFO : aggregate_evaluate: received 2 results and 0 failures
INFO :
INFO : [ROUND 2]
INFO : configure_fit: strategy sampled 2 clients (out of 2)
INFO : aggregate_fit: received 2 results and 0 failures
INFO : ROUND 2 | Client 1 (cid=16776d4558e54926970861e8e939a04d) | Divergence vs mean: 0.027
INFO : ROUND 2 | Client 2 (cid=065534176d4a4f60894ca540cba3ce6b) | Divergence vs mean: 0.027
INFO : configure_evaluate: strategy sampled 2 clients (out of 2)
INFO : aggregate_evaluate: received 2 results and 0 failures
INFO :
INFO : [ROUND 3]
INFO : configure_fit: strategy sampled 2 clients (out of 2)
INFO : aggregate_fit: received 2 results and 0 failures
INFO : ROUND 3 | Client 1 (cid=16776d4558e54926970861e8e939a04d) | Divergence vs mean: 0.029
INFO : ROUND 3 | Client 2 (cid=065534176d4a4f60894ca540cba3ce6b) | Divergence vs mean: 0.029
INFO : configure_evaluate: strategy sampled 2 clients (out of 2)
INFO : aggregate_evaluate: received 2 results and 0 failures
INFO :
INFO : [SUMMARY]
INFO : Run finished 3 round(s) in 297.48s
INFO : History (loss, distributed):
INFO : round 1: 0.6876766395568847
INFO : round 2: 0.5643207287788391
INFO : round 3: 0.5581887340545655
INFO : History (metrics, distributed, evaluate):
INFO : {'bce': [(1, 0.6876766506830851),
INFO : (2, 0.5643207287788391),
INFO : (3, 0.5581887221336365)]}
INFO :
root@d30160989a9d:/srv/ai4os-federated-server/fedserver#
  
```



INFRAEOSC projects synergies:

AI4EOSC, EOSC TITAN and EOSC SIESTA

Federated learning training



Funded by
the European Union



- Sustainability: SIESTA and TITAN were able to lean on AI4EOSC legacy. New INFRAEOSC projects should lean on previous projects,
 - Actively sustaining the goal of a federated EOSC infrastructure.
- Reusing existing architecture and services rather than duplicating efforts
 - Supporting the EOSC goal to improve the efficiency and impact of public investments in research.
- This collaboration is possible due to open source software, but also due to open APIs and interfaces



- This (small) use case is a practical demo of the EOSC acting as a distributed "system of systems"
 - Connecting independent data and services across borders
- Computation is brought directly to the data location, allowing global AI models to be trained without centralising data
- Data owners retain privacy and data control
 - Supporting the EOSC goal of increased data sovereignty
- Cross-border, cross-project federated AI services
 - Supporting EOSC strategic objectives to develop machine-actionable and interoperable research tool



Thank you for your attention

Álvaro López García and Judith Sáinz-Pardo Díaz
on behalf of the EOSC SIESTA and AI4EOSC consortiums
aloga@ifca.es sainzpardo@ifca.es

Advanced Computing and e-Science group
Institute of Physics of Cantabria (IFCA) CSIC-UC



Funded by
the European Union

