LUMI AI Factory, EOSC, Data Labs and Data Spaces for Science

Sara Garavelli
Strategic European Engagement and Coordination Development Manager, CSC-IT Center for Science
sara.garavelli@csc.fi

The ecosystem and the interconnection challenge

- **AI-driven science needs: world-class compute, trusted data, agentic AI software platforms, AI expertise and services.**

- Europe is building these capabilities through AI Factories, Data Labs, Data Spaces, and EOSC.

- The next challenge is integration: **connecting capabilities, stakeholders, governance, and resources** into a functioning **European ecosystem**.
  - **This requires stronger alignment between data and compute initiatives**







Compute is a key asset

- Europe provides multiple pathways to access to **world-class compute at national and European level.**

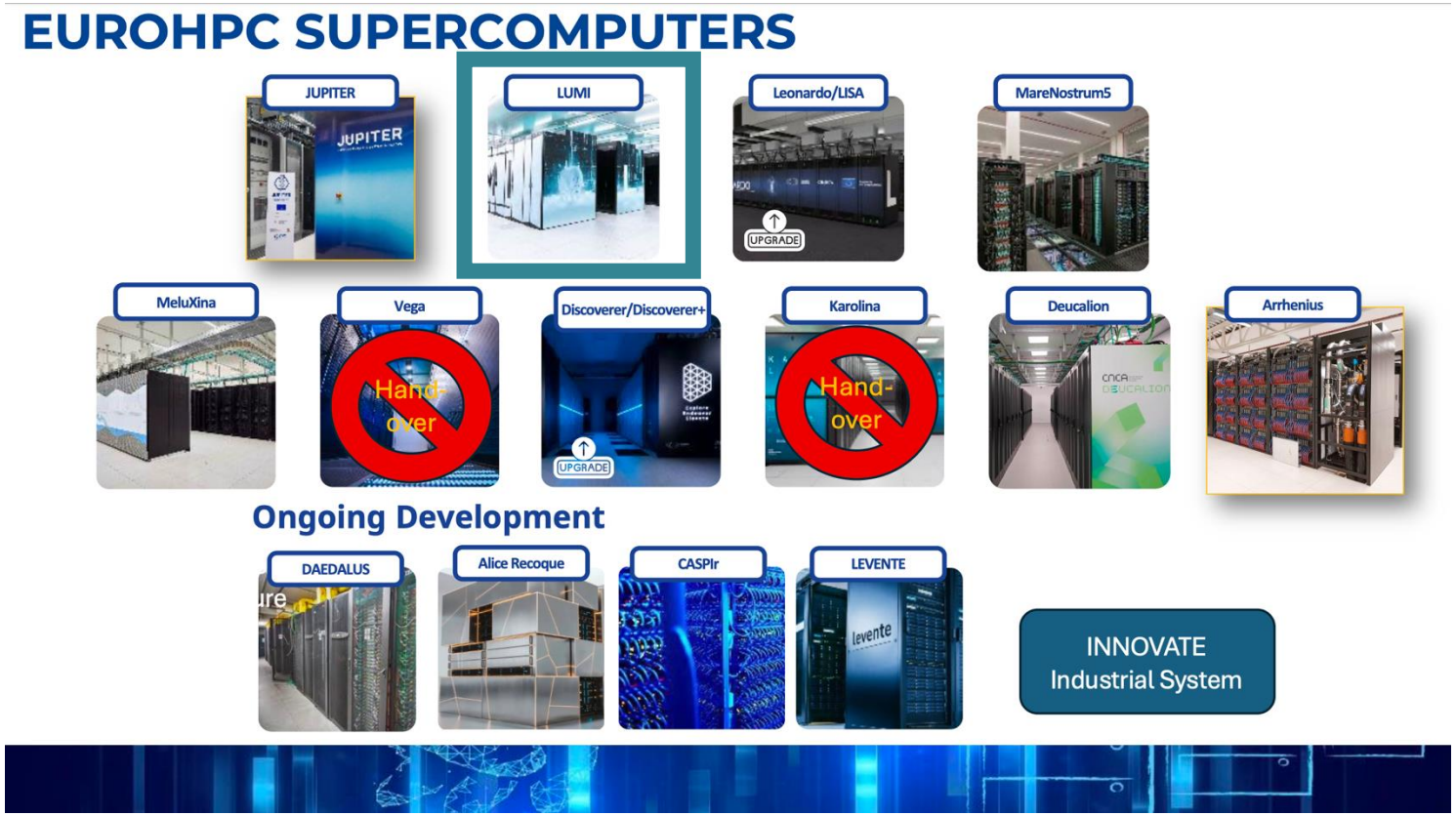
The societal impact of high-performance computing is significant

An investment of €1 into CSC's high-performance computing services generated a

€25-37

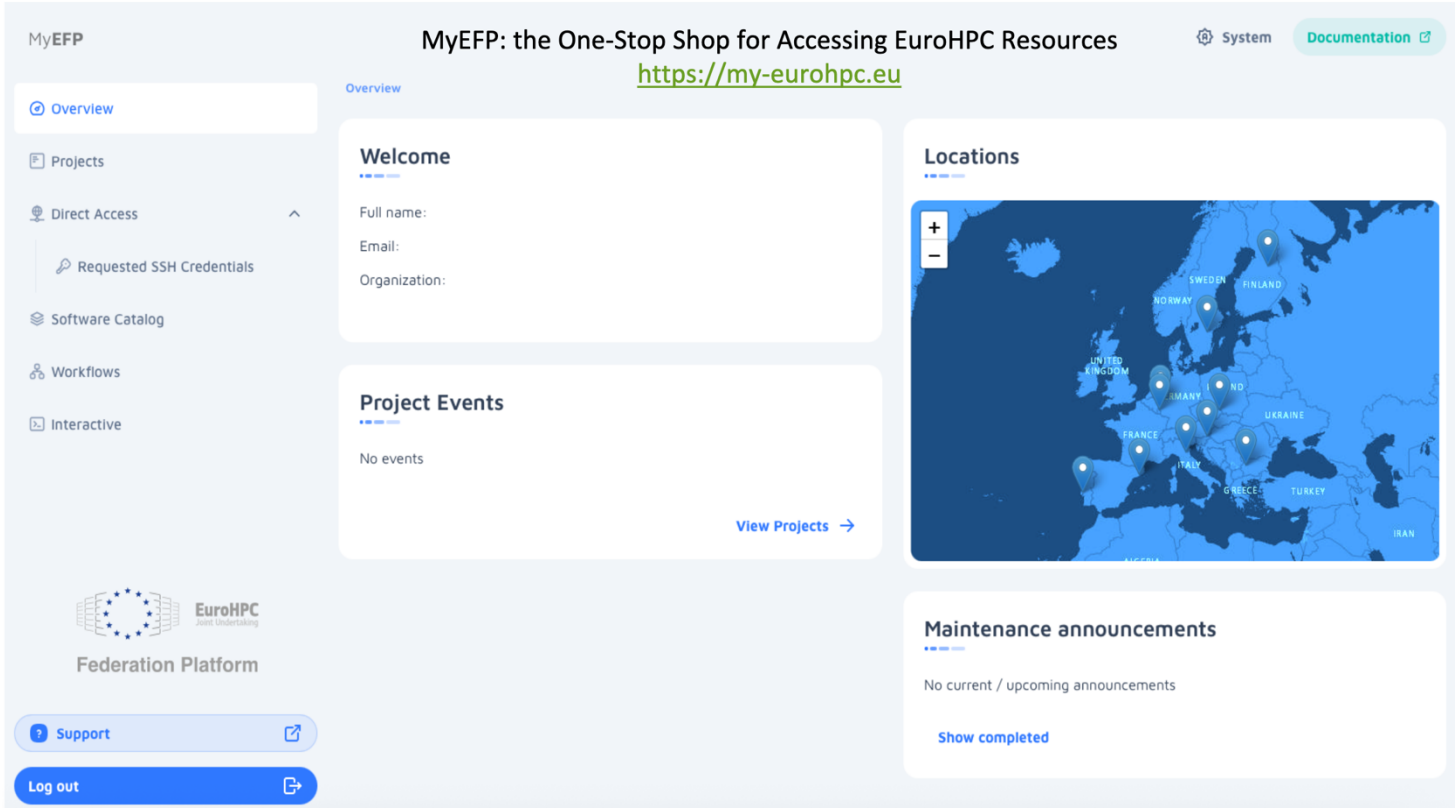
economic benefit to society.

Source: SROI (Societal Return of Investment) analysis conducted by Taloustutkimus in 2024



EuroFP: EuroHPC Federation Platform

- A platform and federation framework which provides a streamlined, **single point of onboarding and** via which user can **access** all systems and services available to them in the EuroHPC ecosystem
- Designed **to support an expanding range of workflows** and enable innovative approaches to complex challenges.



Compute alone does not deliver scientific breakthroughs

To make good use of the capacity, we need expertise and good data

AI Factories and Antennas

AI FACTORIES

BUILDING AN AI ECOSYSTEM IN EUROPE

€2.6 billion investment

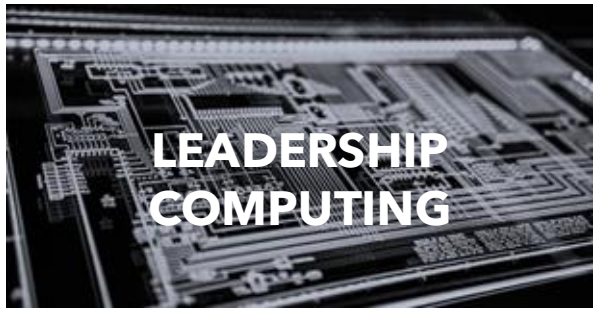
19 AI Factories in 16 European countries

13 AI Factory **Antennas**



LUMI AI Factory

LUMI AIF service offering



Largest AI supercomputers in Europe

- Different sized packages for AI computing
- Supercomputing environment & Kubernetes
- Scalable AI training up to foundation model range



AI expertise from a consortium of 200+ specialists

- Consultation for all phases of AI development
- Special skills in large scale AI and code optimization
- Tailored training in AI and high-performance computing



Access to high value data (Dataset-as-a-Service)

- Integration of data streams
- Confidential and secure data environment
- LUMI hosting is ISO 27001 certified

Explore our services at https://lumi-ai-factory.eu/service-catalogue/

Images: Adi Goldstein, Jaffer Nizami, and John all on Unsplash

Europe’s data is our untapped treasure

- Europe has vast and high-quality research data, but it remains fragmented and hard to use.
- Initiatives like **EOSC** and **Data Labs** are driving the development of a trusted, sovereign and interoperable research data ecosystem across Europe.



World-class compute needs world-class data

– and the infrastructure to connect them.

Data must remain discoverable, reusable, governed and trusted over the long term.

EOSC Federation

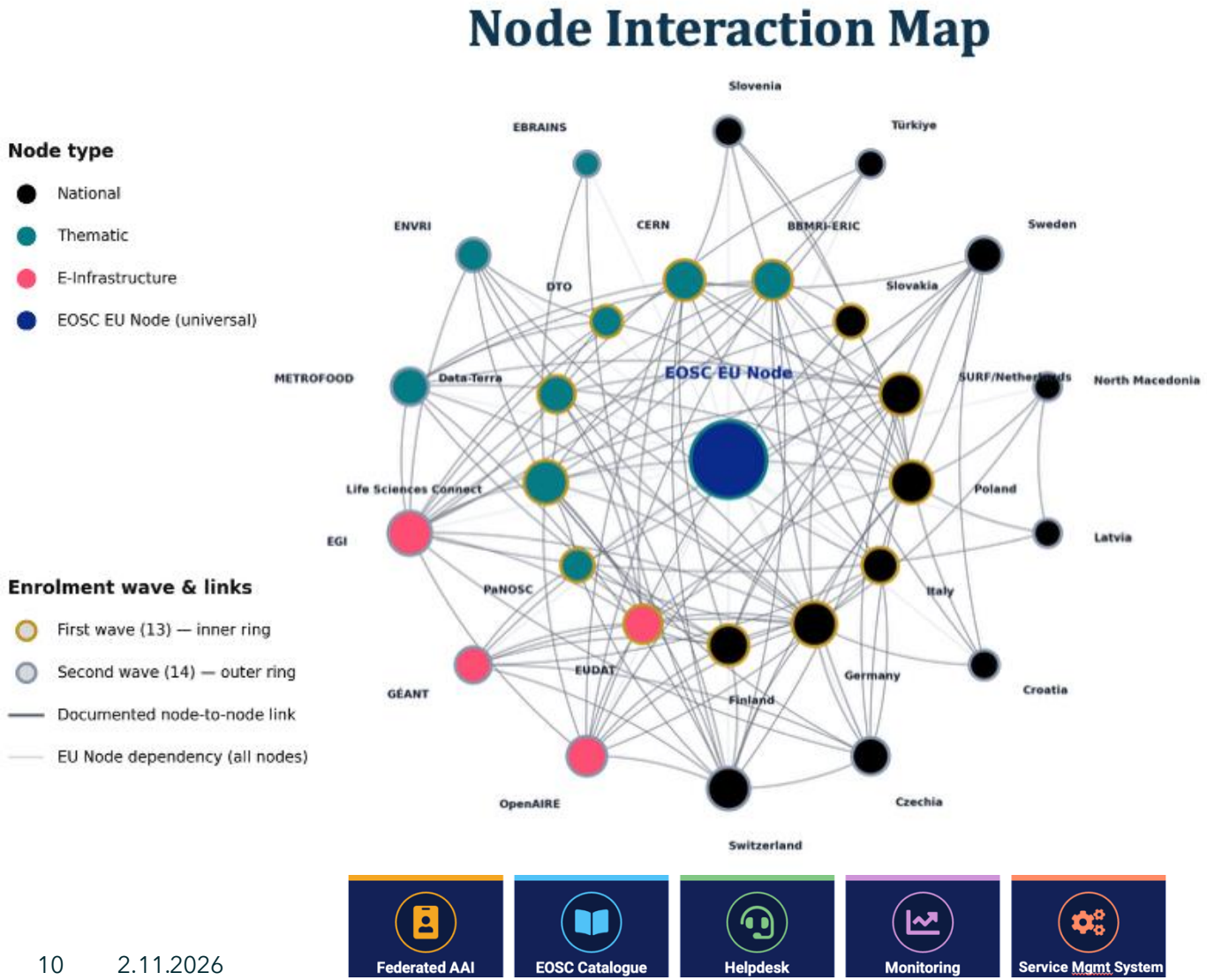
Europe´s federated ecosystem for research data and services

- **Connects existing investments:** EOSC does not build a new infrastructure from scratch. It federates existing national and European research infrastructures (Nodes).
- **Makes research resources interoperable:** Common standards and coordination allow data and services to work together across countries and disciplines.
- **Enables Europe's digital research ecosystem:** EOSC links with initiatives such as EuroHPC and can support emerging AI Factories through access to high-quality research data.

From a single portal to a trusted and sovereign data ecosystem

From a centralised EOSC Portal **to a distributed EOSC Federation in production**

EOSC Federation



- **Shared access to:**
  - Research Data
  - Research Services
  - Resources such as compute and storage
  - Scientific tools
- **Benefits for users:**
  - Find data across borders
  - Reuse trusted datasets, repositories and services
  - Collaborate more easily across borders
  - Accelerate discovery and innovation

eosc 28 nodes, 14 in production

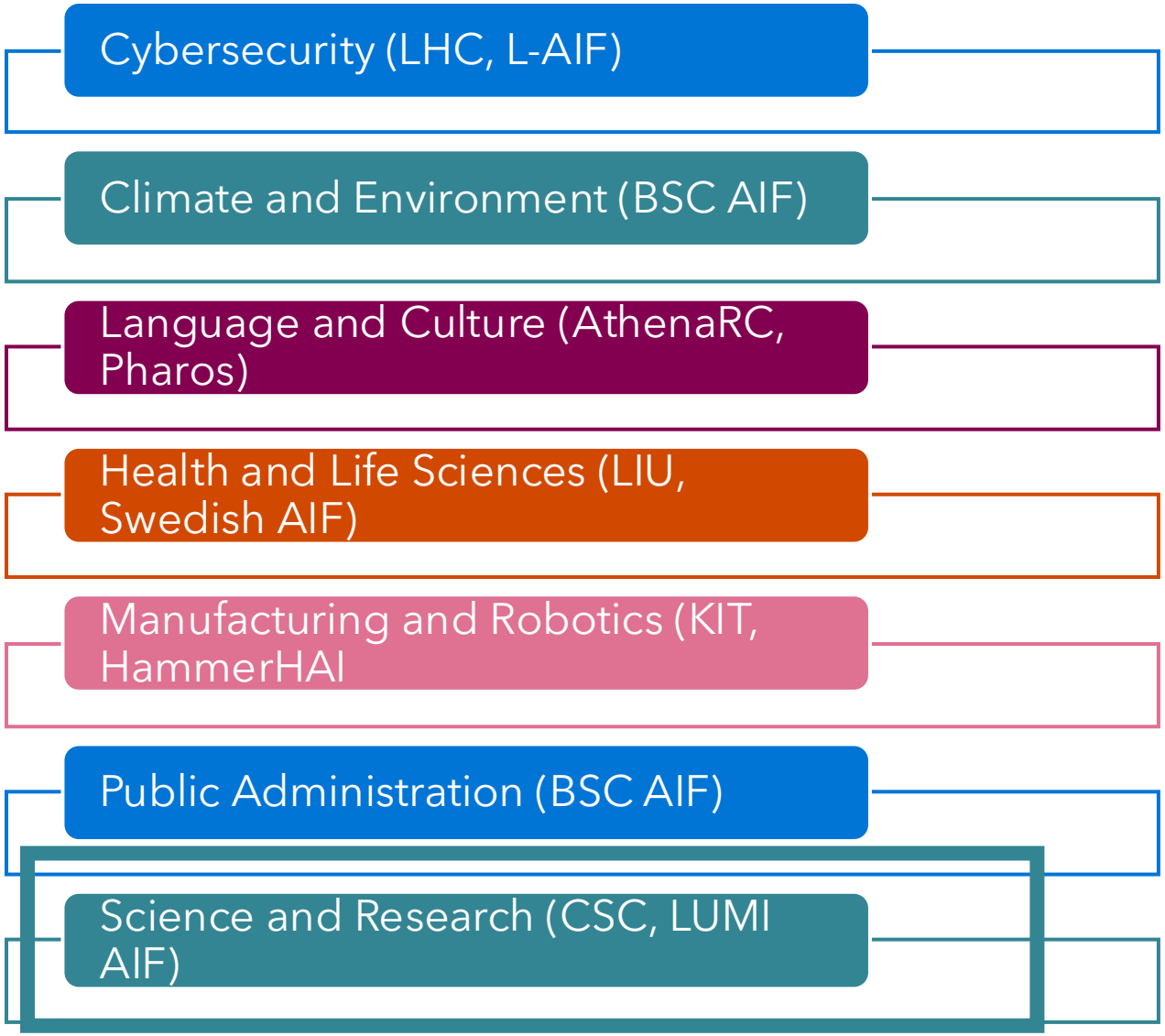


218 EU research organisations currently engaged

A new wave of nodes will be enrolled in 2027

AI Factories (AIFs) Data Labs

- Publication of the Data Union Strategy, November 2025
- Request from EuroHPC JU to start working on the set up of 7 sectoral Data Labs in early 2026
- Appointment of Data Labs coordinators by the EuroHPC JU & start of WGs
  - WGs participants: AIFs and Antennas + other ad hoc experts



AIFs Data Labs: CoopAIFA project

- EuroHPC JU call " Cooperation of AIFs and Factories Antennas" (HORIZON-JU-EUROHPC-2026-COAIF-03), June 2026.
- Type of Action: Research and Innovation Action (RIA)
- Duration: 3 years (Start: 1 November 2026)
- Budget: 25 M€ (50% funding)
- Coordinator: CSC (FI)
- 56 Partners covering all the AIFs and Antennas
- 30 countries represented (FI, NO, ES, EST, DE, CZ, FR, AT, IE, IS, LV, BE, BU, MT, PL, MD, CH, HU, IT, LUX, SE, LT, NL, GR, CY, RO, RS, SK, UK, MK)

Main Focus

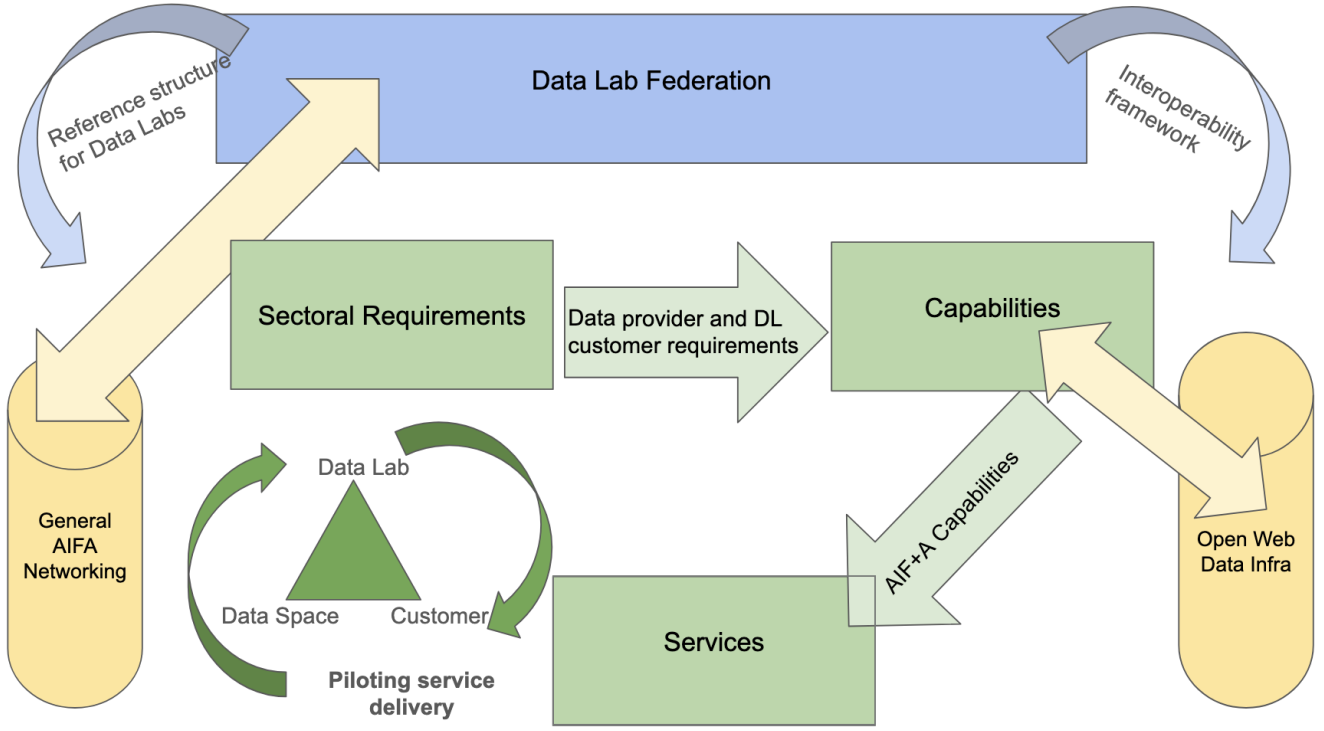
- **Coordinating and networking of AI Factories and Antennas**
- **Networking of AIF Data Labs** to enable the secure and efficient sharing of data and data services between different AIFs & Antennas. This will be made possible through the use of SimpI, which will facilitate the exchange of data and ensure compliance with EU regulations and standards.
- **Provision of EU open web data**

Data Labs in the AIFA collaboration project

To **establish and operationalise a federated Data Lab service layer that systematically connects AIFs+As, Common European Data Spaces, and AI developers**.

The CoopAIFA project will create a **blueprint for AIFA Data Labs**, establish governance structures, and implement an interoperability framework.

Sectoral Data Labs will define **requirements for services through real-life use cases.**



Data Lab capabilities/services:

- Metadata Management Discovery and Data Ingestion Pipelines
- Interoperability building blocks
- Secure processing
- AI-ready FAIR reference datasets
- Data-related service discovery
- Compliance and cybersecurity
- Synthetic Data Generation and Validation Service Framework

Connecting Europe’s AI and research infrastructure for science:

EuroHPC
+
Data Labs + EOSC
+
AI skills & tools

How does this translate on a country? The Finnish example

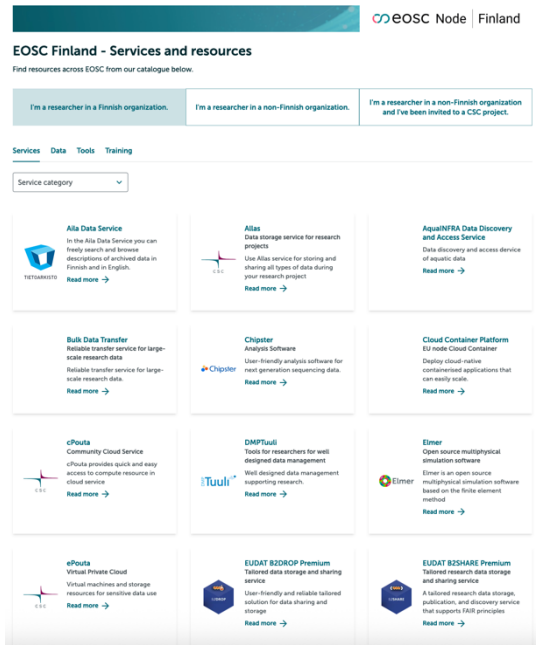
Connecting Europe's AI and research infrastructure for science

EuroHPC



World-class compute
LUMI AI

EOSC



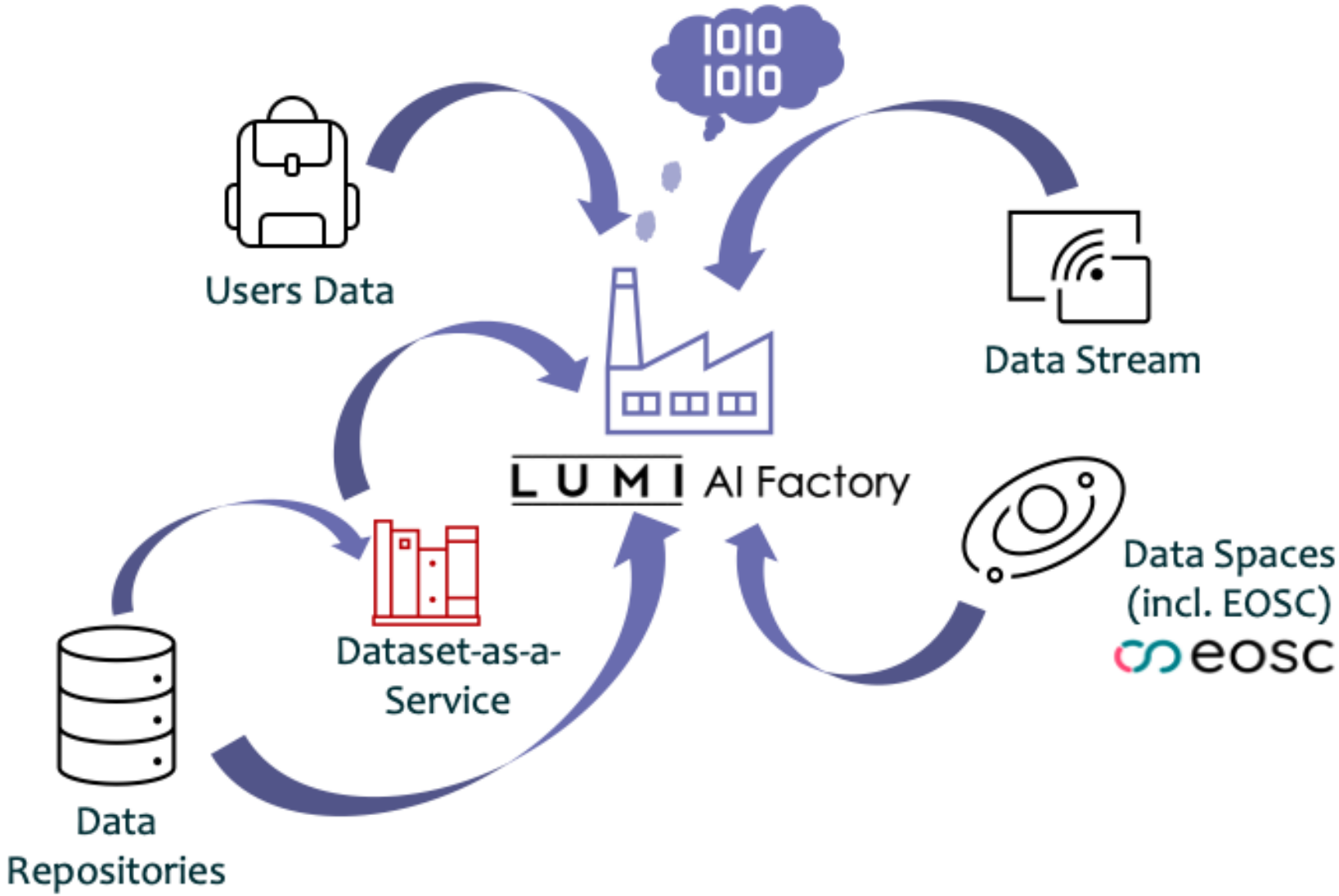
FAIR Research Data
Research Services
Scientific tools

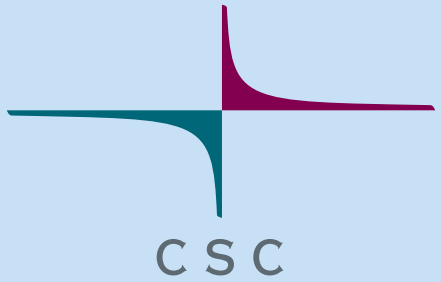
LUMI AIF



AI expertise & consultancy
AI-ready FAIR data
Data and models as a service

Our future vision





Thank you

sara.garavelli@csc.fi

Follow us

LinkedIn

Instagram

Facebook

YouTube

csc.fi