ISGC 2026

EGI's Path Towards an AI-ready Compute and Data Continuum for Research

Sergio Andreozzi
Head of Strategy, Innovation and Communications, EGI Foundation (NL)
sergio.andreozzi@egi.eu

Outline

1. **AI Reshaping Science**
   AI as a scientific capability, the rise of agentic systems, and the infrastructure challenge

2. **Europe's Policy Response**
   AI Continent Action Plan, RAISE, and the SCIANCE project

3. **EGI: Towards an AI-Ready Compute and Data Continuum**
   EGI Federation, new 5-year strategy, AI-related Projects and Services

AI Reshaping Science

AI as a Scientific Capability

SEE
Perception & Analysis

Process vast and complex data, finding meaningful patterns that would be impossible for humans to detect

e.g. analysing detector images, identifying faint signals in noisy data

DECIDE
Reasoning & Optimisation

Evaluate complex scenarios to optimise decisions and experimental strategies

e.g. classifying which collisions to save, optimising telescope scheduling

CREATE
Generation & Simulation

Generate new data, models, or solutions that accelerate the pace of research

e.g. ultra-fast simulations reducing hours to seconds, AI-assisted scientific coding

From Proteins to Gravitational Waves

AlphaFold

Predicted 200M+ protein structures, transforming life sciences and drug discovery. Built on EMBL-EBI open data.

Deep Learning | Neural Network Architectures

ECMWF Weather AI

European Centre's ML models now rival traditional physics-based simulations for medium-range weather forecasting

Graph Neural Networks | Ensemble ML

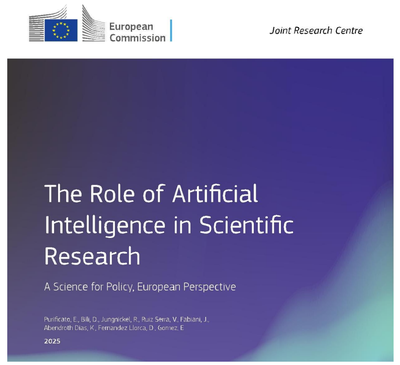
LIGO / Virgo

AI detects gravitational wave signals buried in detector noise. Europe's Virgo (EGO) collaborates across continents in a federated model

CNNs | Matched Filtering | Signal Processing

The Infrastructure Challenge

- Large-scale GPU compute, curated scientific datasets, interoperable platforms
- Growing demand that no single institution or country can address alone
- Need for federated, sovereign, open-source infrastructure at continental scale



The Rise of Agentic Science

Multi-agent AI systems are evolving from tools into autonomous research partners capable of hypothesis generation, experimental design, and validation

The AI Scientist v2

Sakana AI

Single Agent | End-to-End Orchestrated

Uses agentic tree search to autonomously generate, execute, and write up ML experiments. First AI-authored paper to pass peer review.

arxiv.org/abs/2504.08066

AI Co-Scientist

Google DeepMind

Multi-Agent | Centrally Coordinated

Agents “generate, debate, evolve” approach to hypothesis generation; applied to biomedical research, with validated in-silico discoveries.

arxiv.org/abs/2502.18864

ScienceClaw × Infinite

MIT — March 2026

Decentralised Swarms | Self-Organising

300+ composable skills, no central coordinator. Agents self-organise around a global need index. Open-source, provenance-linked discovery.

github.com/lamm-mit/ScienceClaw

Why Infrastructure Matters

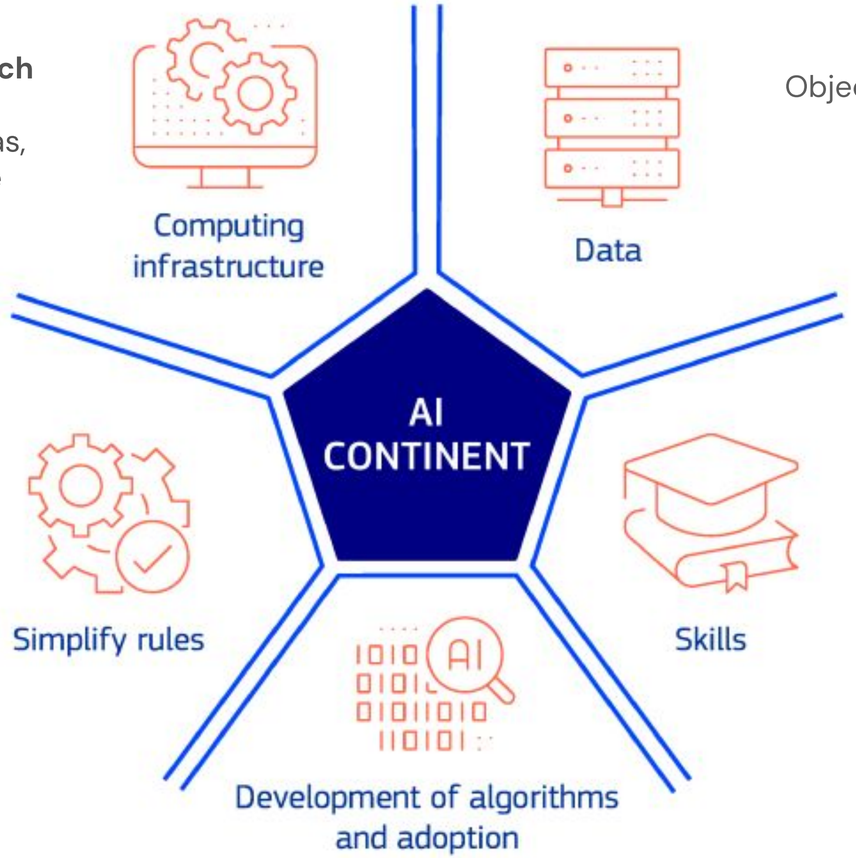
These systems demand infrastructure that can scale, be trusted, and remain sustainable

Europe's Policy Response

- EU AI Continent Action Plan
- RAISE: The Resource for AI Science in Europe
- Sciance: AI in Science Coordination Project

EU AI Continent Action Plan

Objective: **Train and develop AI models, boost research in sustainable infrastructures**
Mechanism: 19 AI Factories selected, AI factory antennas, future AI Gigafactories (>100K chips), triple data centre capacity



Objective: **Improve data access for businesses and public admins**
Mechanism: Data Union Strategy; Data Labs inside AI Factories to curate high-quality datasets

Objective: **Boost citizens' trust and provide legal certainty**
Mechanism: AI Act Service Desk with free customised tools and advice for businesses

Objective: **Expand AI competencies and attract global talent**
Mechanism: AI Skills Academy, international fellowships, pilot generative AI degree

Objective: **Deploy AI in strategic sectors: healthcare, automotive, advanced manufacturing**
Mechanism: Apply AI Strategy with direct support to businesses and public administrations

Source:AI Continent Action Plan COM(2025)165 - 16 April 2025
https://digital-strategy.ec.europa.eu/en/library/ai-continent-action-plan

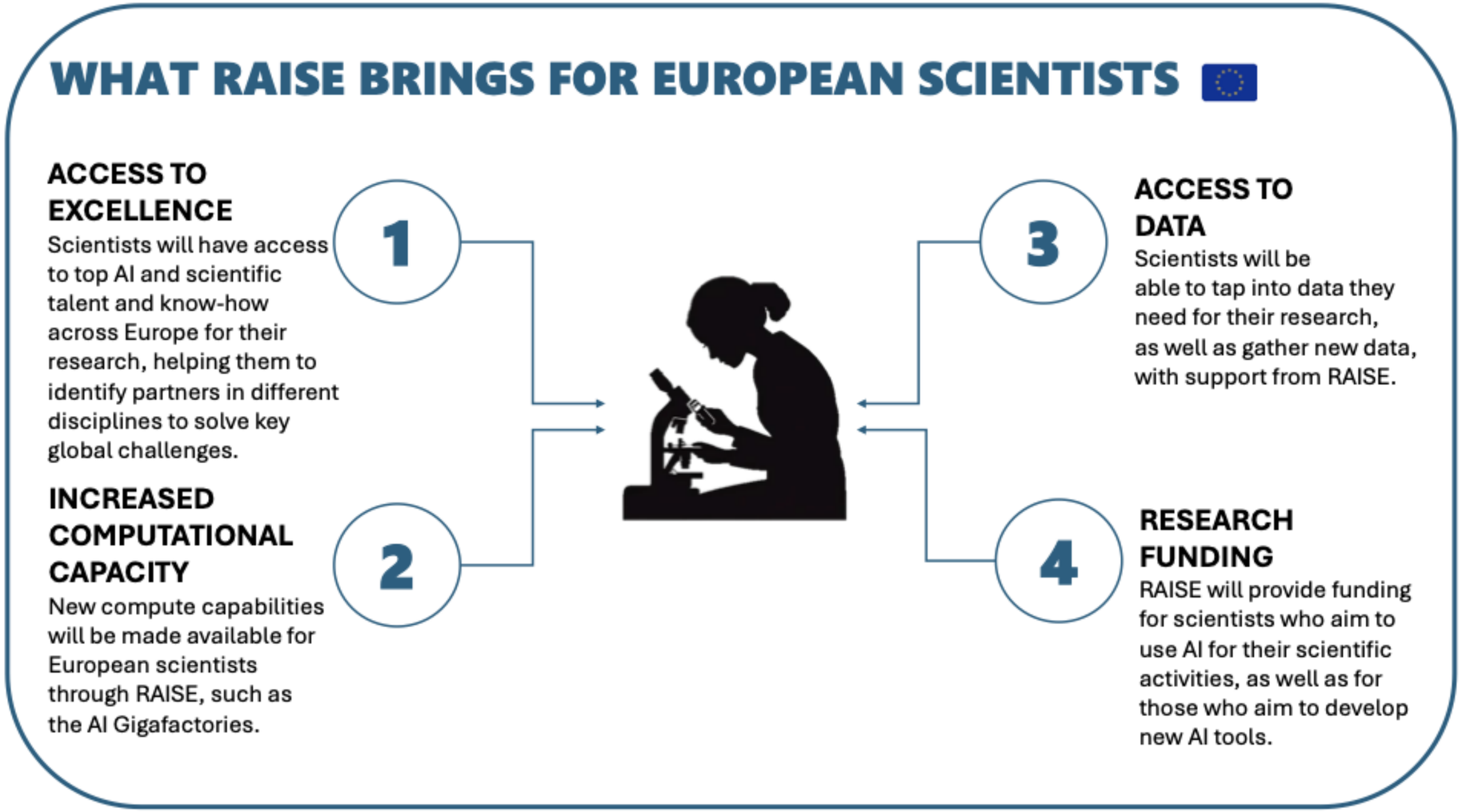
RAISE: The Resource for AI Science in Europe

What is it?

A virtual European institute that pools, aligns and coordinates essential AI resources.

Part of the **Apply AI Strategy** and the **European Strategy for Artificial Intelligence (AI) in Science**

Launched in Nov 2025



Source: RAISE – Resource for AI Science in Europe
https://research-and-innovation.ec.europa.eu/strategy/strategy-research-and-innovation/our-digital-future/european-ai-science-strategy/raise-resource-ai-science-europe_en

EC Main RAISE Actions Roadmap

Horizon Europe (WP25)

- Facilitated cooperation for AI in Science (RAISE secretariat) (Q4/25) (EUR 3 Million)
- AI Foundation Models in Science (materials science, climate change, pollution & agriculture) (Q1/26) (EUR 30 Million)

Coming soon: Horizon Europe (WP26-27)

- Thematic Networks of Excellence for materials science and agriculture & environment science (EUR 28 Million)
- Pilot of the "Science for AI Pillar" of RAISE (EUR 17 Million)
- RAISE Doctoral Networks (EUR 30 Million)
- Scientific Laboratory Automation (EUR 32 Million)

Other Ongoing Actions

- Horizon Europe contribution to the AI Gigafactories (EUR 600 Million)
- ERA Action on AI in Science
- RAISE Scientific Advisory Board (Q1-2/26)

Future Horizon Europe (FP10) / Next MFF

- Proposed Moonshot on next generation AI
- Robust & durable anchoring of RAISE

SCIANCE AI in Science Coordination project (RAISE)

Map the European AI in Science Landscape

Evidence baseline with cross-disciplinary and thematic reviews

Co-Define Research & Innovation Priorities

Co-creation workshops, AI in Science challenges, interdisciplinary cooperation

Assess Infrastructure Scenarios

Identify gaps in computing, data, governance, and skills; co-design upgrade scenarios

Enable Community Building & Sustainability

Pilot RAISE Secretariat, Digital Hub, Academy; co-organise AIS Summits

SCIANCE feeds Evidence Base, SRIA, and Implementation Roadmap into RAISE

13 partners (including EGI Foundation) | Coordinated by European Science Foundation | https://sciance.eu/

SCIANCE

Scientific Pilot Areas

Physics & Astronomy

Nikhef & University of Manchester

Materials Science

Constructor University

Earth Sciences

Twente University ITC

Life Sciences

Euro-BioImaging

Social Sciences & Humanities

CNR-ISTI (SoBigData)

10 AI in Science Working Groups being established

- Astronomy and Fundamental Physics
- Life Sciences
- Social Sciences and Humanities
- Data Science and Advanced Analytics
- AI and Computing Infrastructures
- Materials Science
- Earth Sciences
- AI Research and Experimental Design
- Open Science and Trustworthy AI
- Industry, Innovation and Frugal AI

Deadline for applications
10 April 2026

EGI: Towards an AI-Ready Compute and Data Continuum

- Brief intro to EGI
- New strategy 2026-2030
- Recent and Next AI-related Projects
- Incubating Initial AI services

EGI Federation_

Digital Infrastructure For Data-Intensive Research

+2,700
Research Publications/Year

+150K
Users from 182 countries

+250
Research projects & collaborations

Participant Country

Associated Participant Country

Participant

Associated Participant

Headquarter of international research organisation or institutional representative

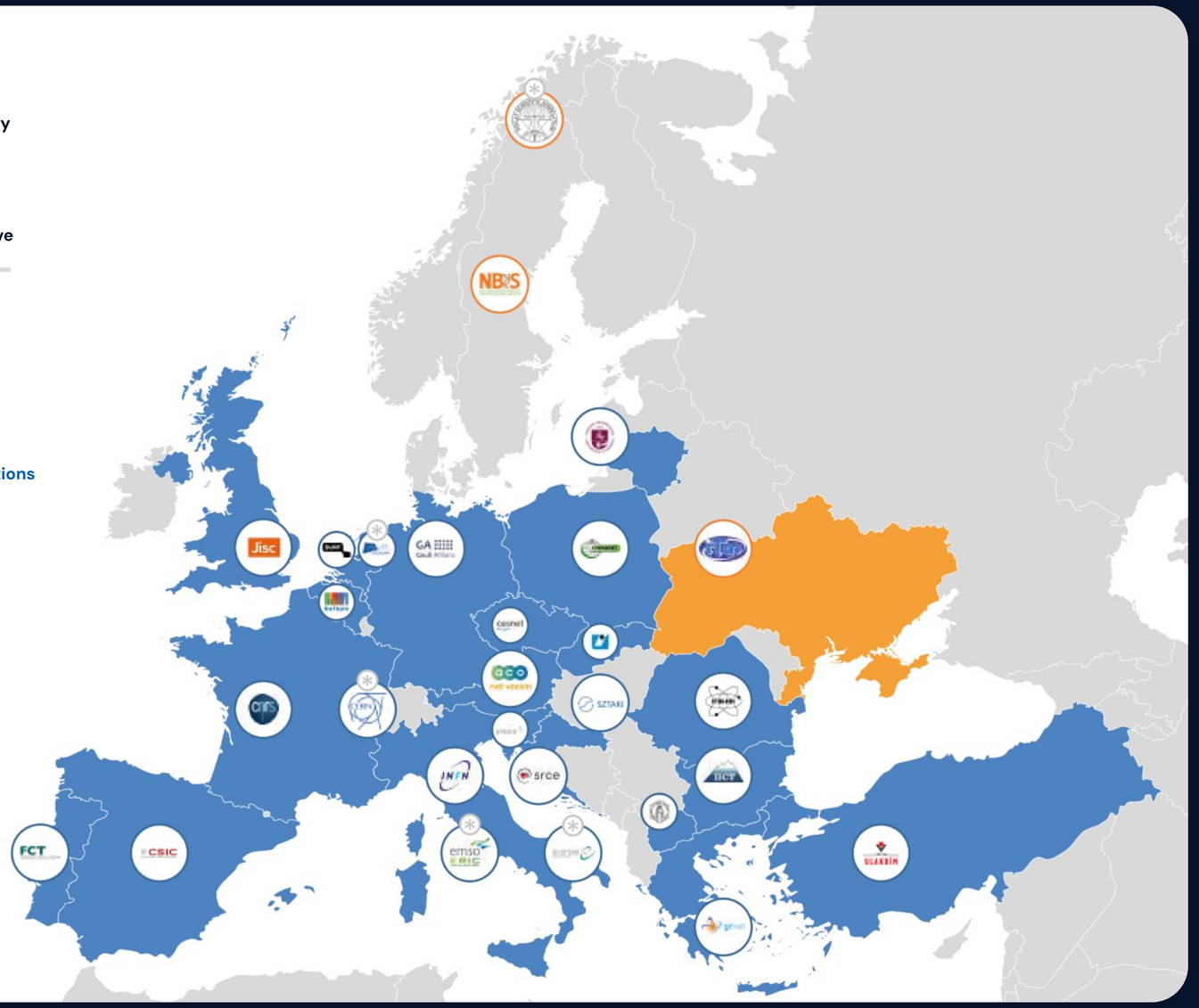
EGI Council Participants:

28

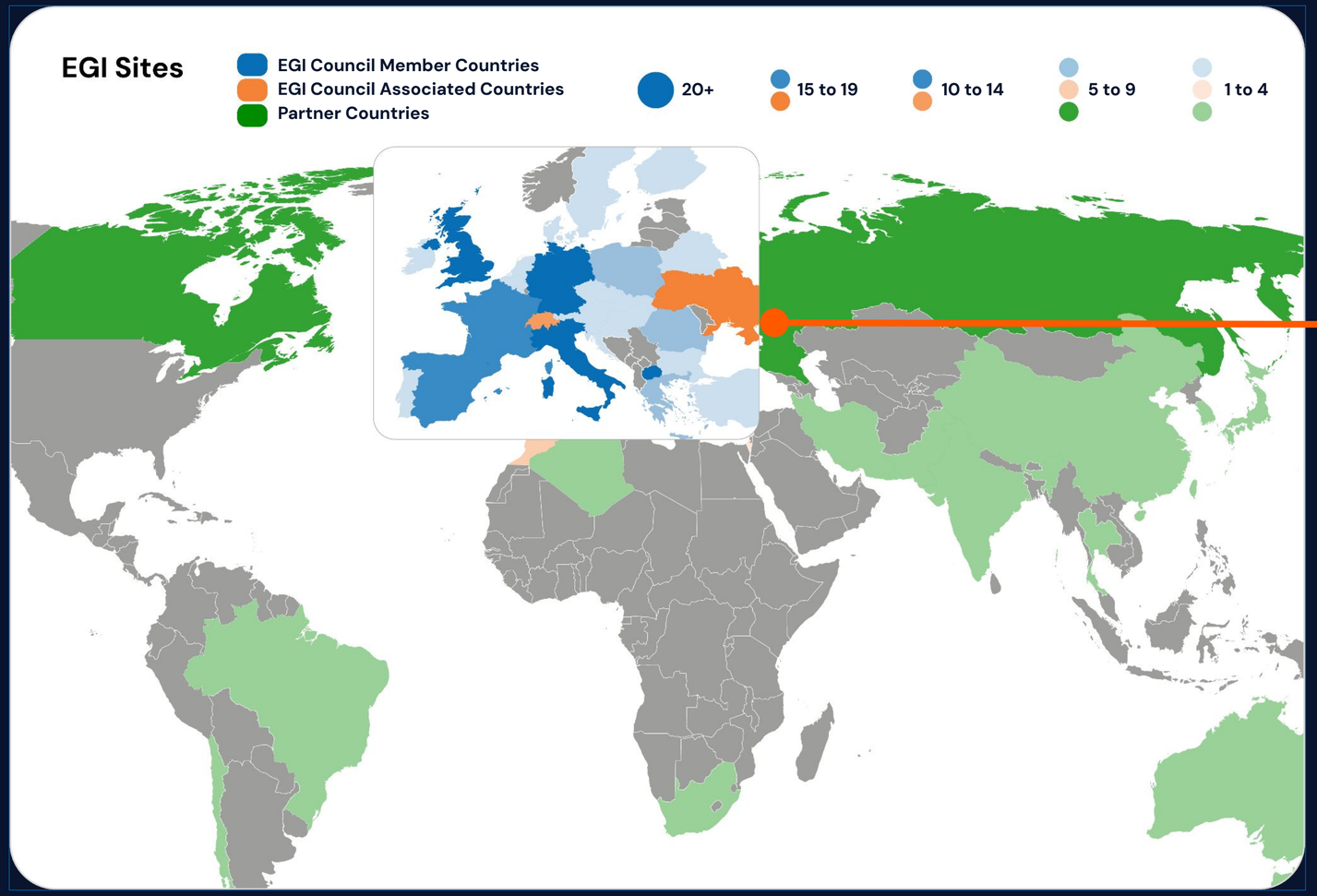
21 countries

5 international research organisations

2 institutional representatives



A Single International Digital Research Infrastructure_



+210 Data centers
42 Countries

Long-term collaboration with Academia Sinica Grid Computing Centre (ASGC)

EGI Services for Research_

Compute

Cloud Compute
Run virtual machines on demand with complete control over computing resources

Cloud Container Compute
Run Docker containers in a lightweight virtualised environment

High-Throughput Compute
Execute thousands of computational tasks to analyse large datasets

Software Distribution
Publish and access software efficiently across multiple sites

Compute Orchestration

Workload Manager
Manage computing workloads in an efficient way

Infrastructure Manager
Use cloud orchestrator to deploy and configure complex virtual infrastructures

Applications

Notebooks
Create interactive documents with live code, visualisations and text

Replay
Reproduce and share research on a notebooks-based platform

Storage & Data

Datahub
Access key scientific datasets in a scalable way

Data Transfer
Transfer large sets of data from one place to another

Online Storage
Store, share and access your files and their metadata on a global scale

Security & Identity

Check-in
Login with your own credentials

Secrets Store
Easily retrieve, manage, and rotate credentials, API keys, and other secrets through their lifecycle

Training

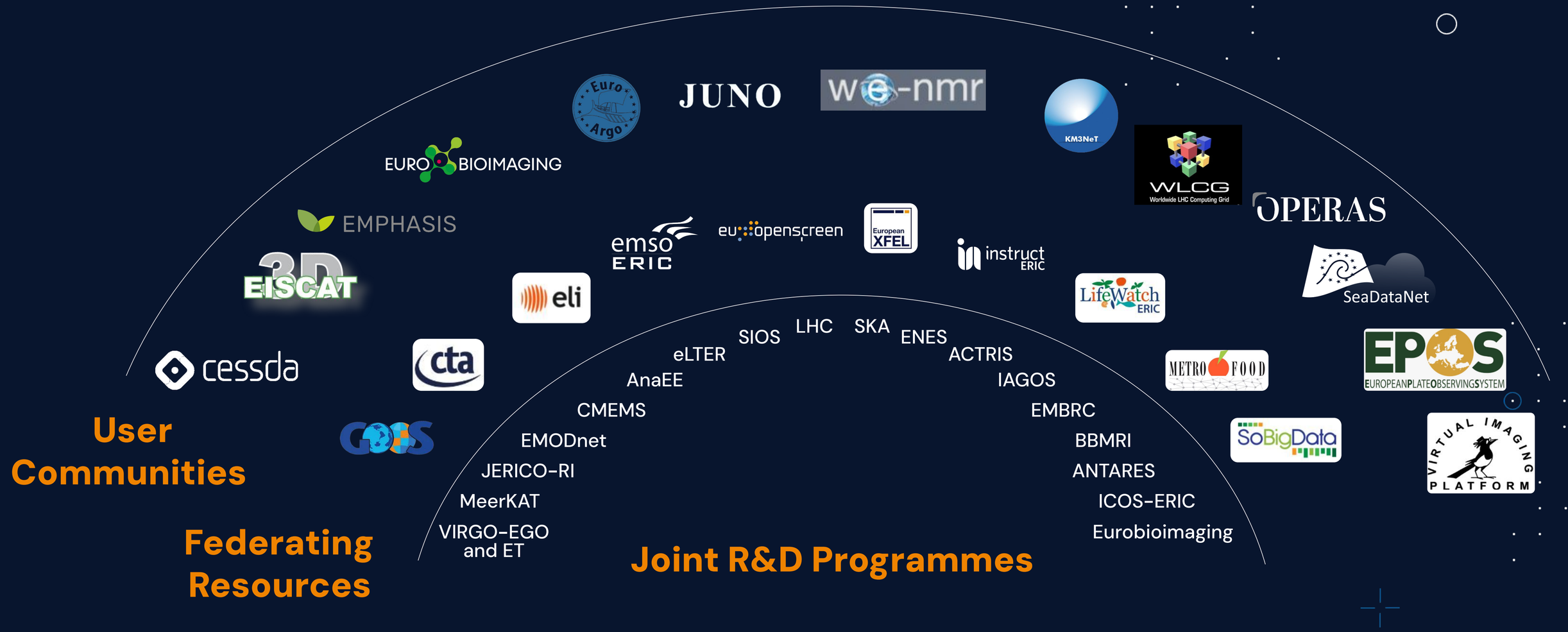
FitSM Training
Learn how to manage IT services with a pragmatic and lightweight standard

ISO 27001 Training
Learn how to manage and secure information assets

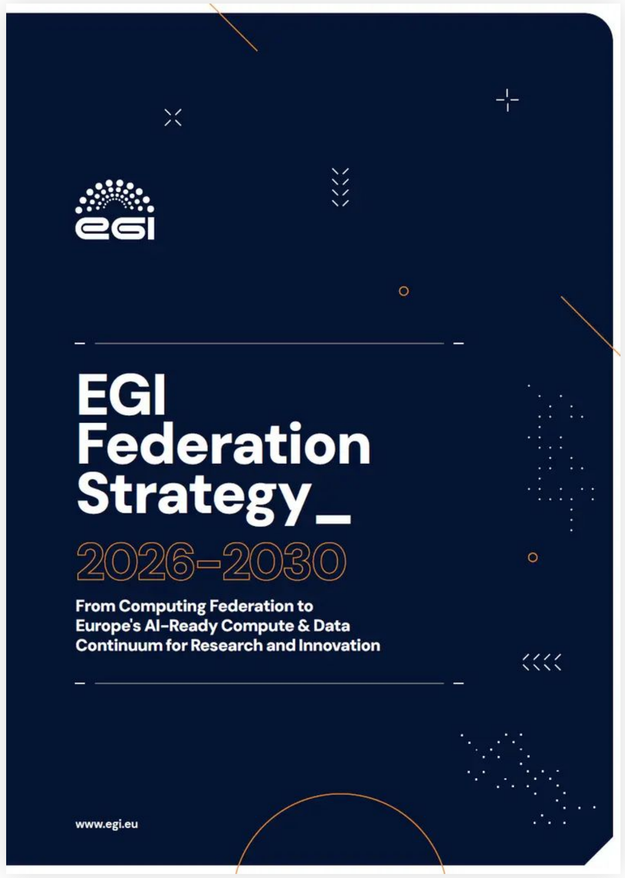
Training Infrastructure
Dedicated computing and storage for training and education

also: technical support, consultancy, research & innovation

Thematic Infrastructures Partnering with EGI



Strategy 2026-2030 at a Glance



https://www.egi.eu/publication/egi-federation-strategy-2030/

Overarching Vision
A future where advanced computing and shared data resources seamlessly empower every field of research to achieve transformative discoveries and address global challenges

Mission
We accelerate the frontier of data-intensive research by federating digital capabilities, fostering collaboration, and driving continuous innovation

Ambition 2030
Europe's essential federated AI-ready compute-data continuum for research communities, where innovation and investment compound for pan-European impact

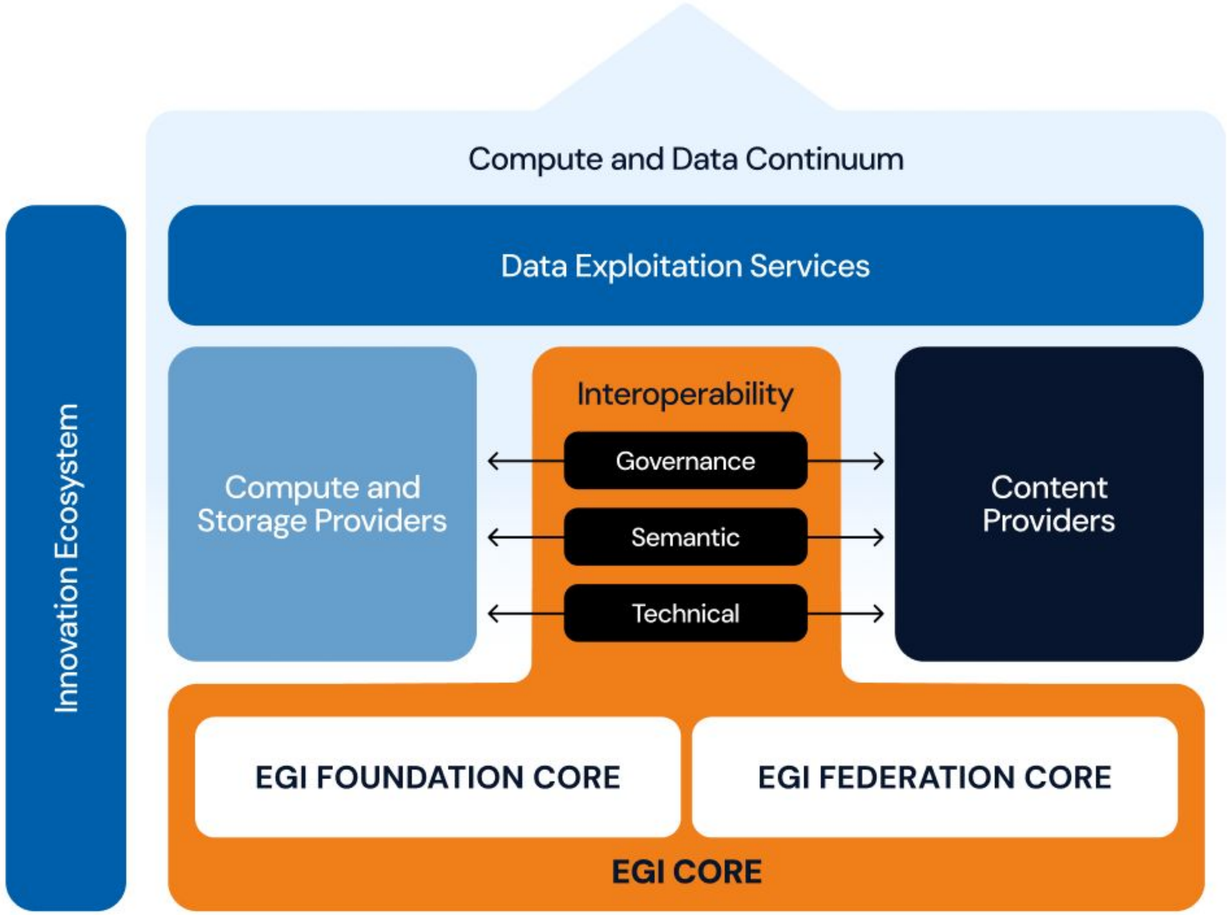
Value Streams

- Powering the Federation
- Enabling Scientific Computing
- Enabling Sustained Innovation

Strategic Pillars

Sovereign Compute-Data Federation	Cross-Border Access Model	Innovation Ecosystem
SG1: Compute-Data Ecosystem	SG3: Sustainable Model	SG5: Key R&I Actor
SG2: AI-Ready Research	SG4: Diversified Funding	SG6: Connected Community

Towards a Compute–Data Continuum

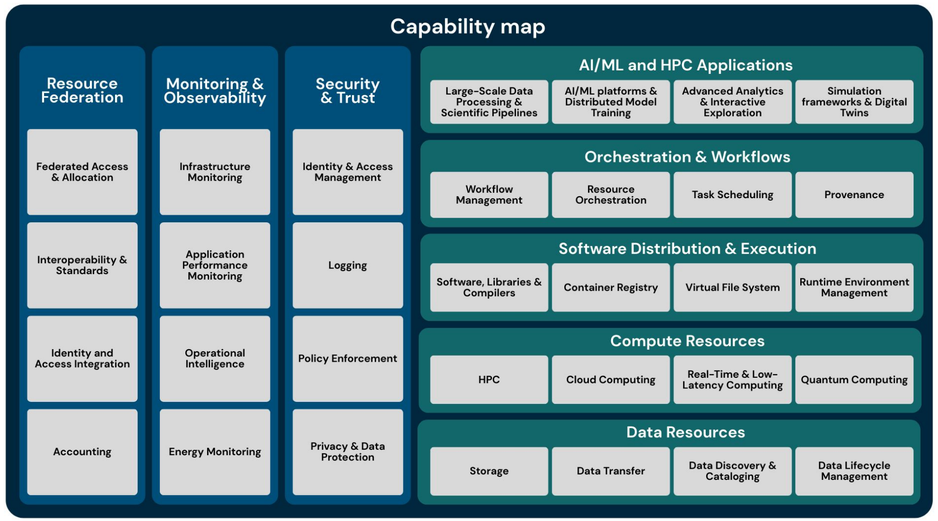


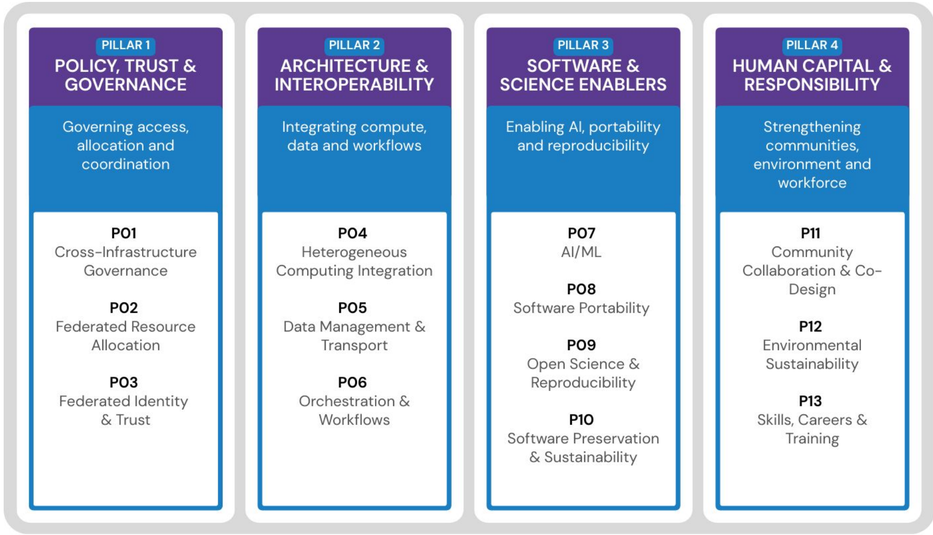
EGI: Towards an AI-Ready Federation

Flagship projects

SPECTRUM

Compute and Data Continuum for Data-Intensive Science





Technical Blueprint

- Capability map with 8 layers
- 8 technical activities with recommended actions
- Co-designed with 14 use cases from 6 Working Groups

Strategic Research, Innovation and Deployment Agenda

- Strategic Research, Innovation & Deployment Agenda
- 13 priorities across 4 pillars

Source: public drafts https://zenodo.org/records/17901519 - more info: spectrumproject.eu

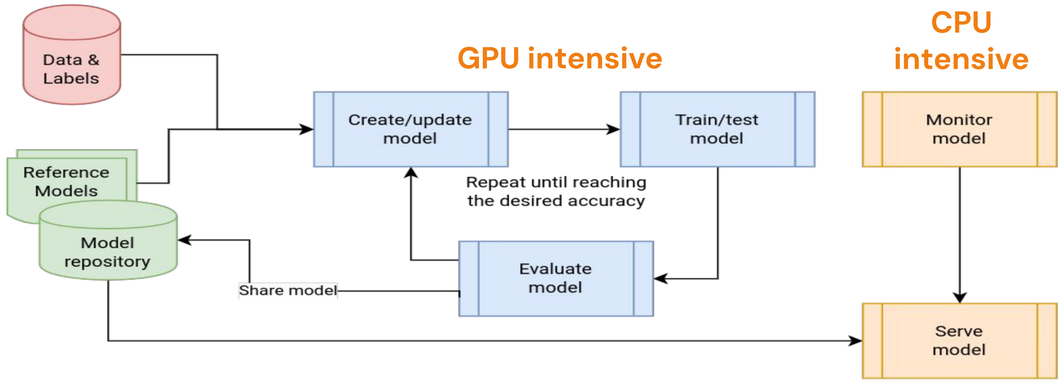
Final event: Brussels, 3 June 2026 https://indico.egi.eu/event/6918/

Building an AI-Ready Federation

Computing	Expand GPU and HPC capacity; deploy Data Exploitation Environments and AI platforms across federated sites
Data	Integrate scientific datasets with FAIR principles and AI-compatible licensing; interoperate with domain data spaces
AI Adoption and Services	Provide AI platforms (AI Lab, FedLLM) and adoption support within scientific disciplines and core services
Skills and Training	Train infrastructure providers to offer AI services; train researchers to adopt AI tools and workflows; develop federated AI competence center
Cross-Border Access	Enable seamless cross-border access through trusted authentication, interoperable services, and simplified policies

iMagine AI for Aquatic and Environmental Science

Thematic Services



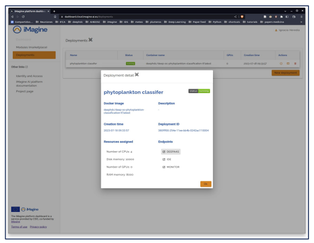
8 internal, 7 external use cases

Benefitting **10+** Research Infrastructures &

EOSC AI4EU

Platform Service





iMagine AI Platform: Generic, scalable platform for developing and sharing AI/ML applications.

AI4 | EOSC
AI technology

Infrastructure

EGI, CSIC Consejo Superior de Investigaciones Científicas, INCD, TÜBİTAK, WALTON Institute for Information and Communication Systems Science

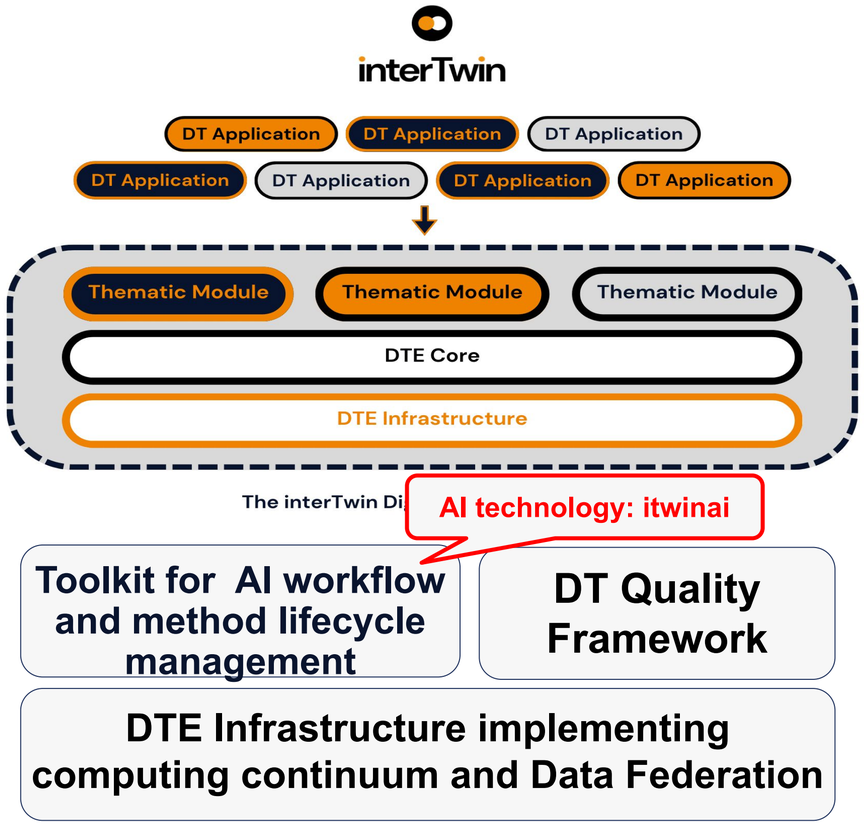
- **1500** TB-months → **843 consumed**
- **132,000** GPU-hours → **365,000 consumed**
- **6 Million** CPU-hours → **11 Million consumed**

4 federated cloud sites

https://www.imagine-ai.eu/

interTwin Digital Twin Engine for Science

interTwin develops a Digital Twin Engine to support the development of Digital Twins for various science domains



Early warning for Extreme events
Deltares, EURAC, TU Wien

WildFire Hazard Map Generation
CMCC, CNRS, Univ. of Trento

Tropical Cyclone Detection
CMCC, CNRS, Univ. of Trento

Extreme events impacts
CERFACS, EURAC, Deltares

Radio Astronomy
Max Planck Society

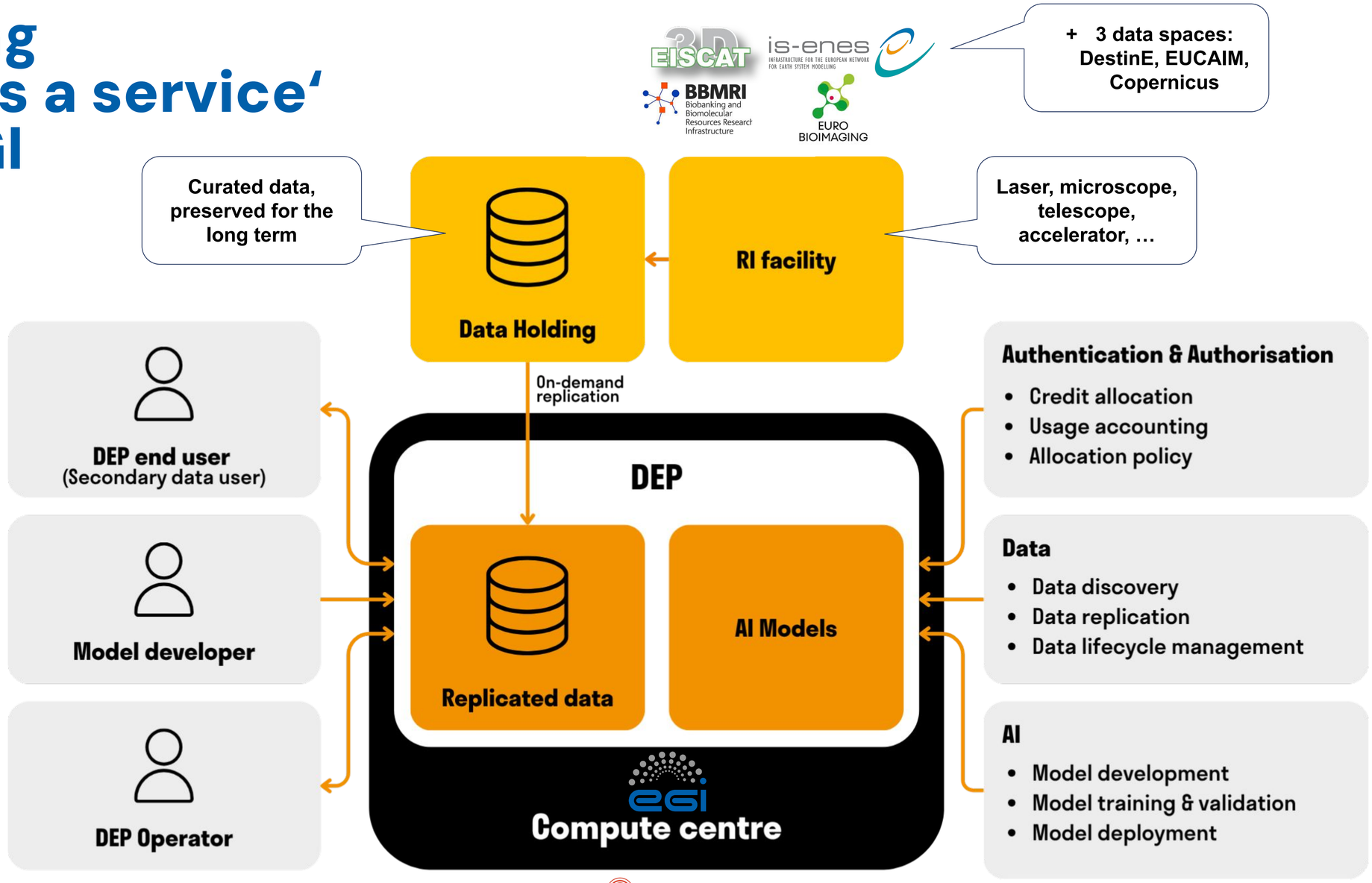
High Energy Physics
CERN, CNRS

Lattice QCD
CSIC, CNRS, ETHZ

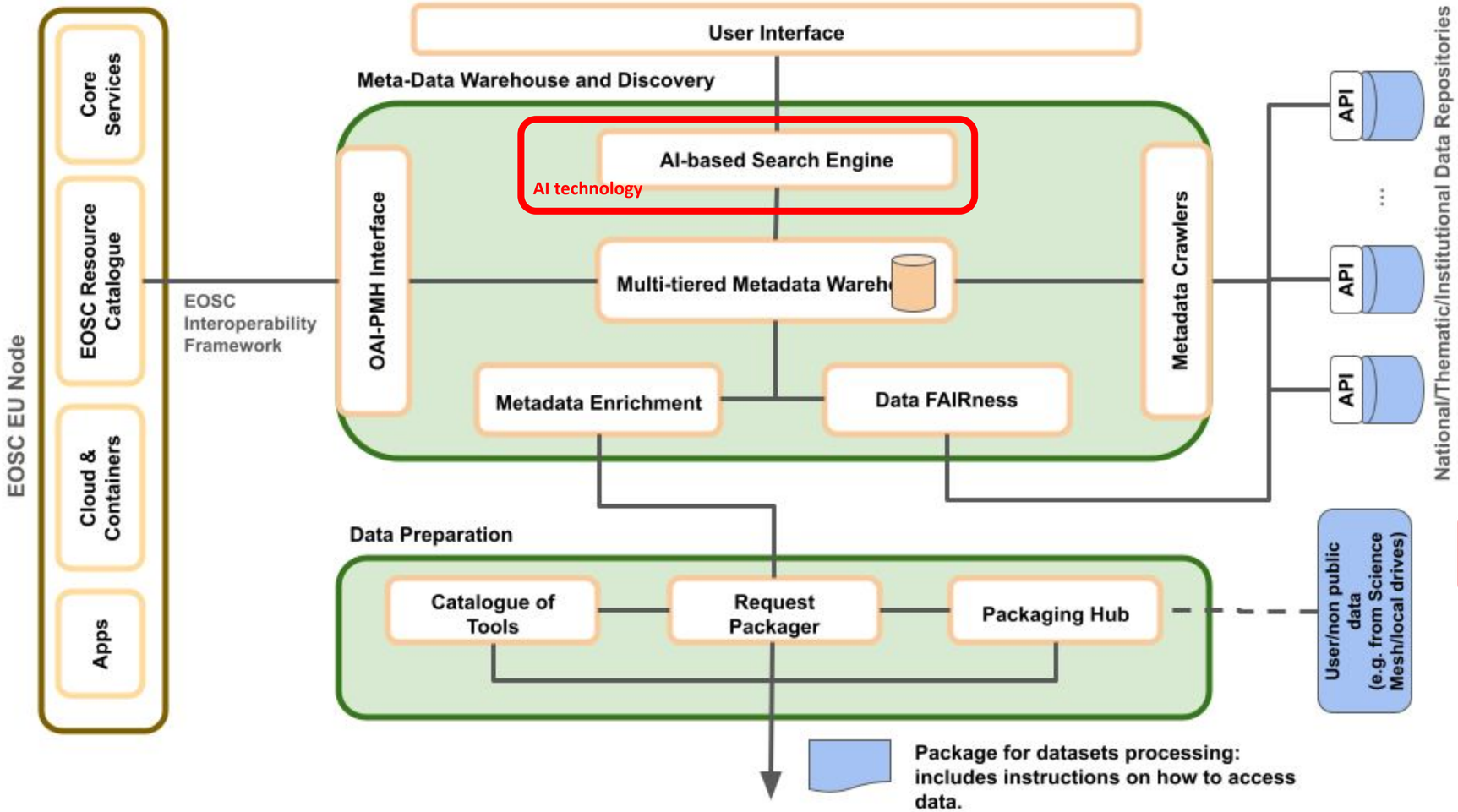
Gravitational Wave - Astrophysics
INFN

Bringing ‘data as a service‘ into EGI

Piloting and validating
- DEP Technology
- Governance policies
- Business model



eosc | DATA COMMONS



The role of AI in EDC:

- Discovery of datasets: LLM powered search over the metadata warehouse and generation of summaries of the search results
- Support the linking of datasets and analytics tools - go beyond simple file type association
- Potential usage for metadata enrichment in the warehouse (TBD)
- Relying on external LLM providers AI technology (e-infra.CZ + Mistral as backup)
- Development of MCP for searching in the warehouse

https://www.eosc-data-commons.eu/

New AI-related Projects Starting in 2026

ENSURE

GenAI digital twins for greening research infrastructures

Data centre digital twins using GenAI and advanced monitoring to model, predict, and reduce environmental footprint. Coordinated by EGI Foundation. GreenDIGIT follow-up.

FLUID-AI

AI4EOSC extension with AI-ready data

Extending FAIR principles for machine-actionable, interoperable, and trustworthy AI/ML assets. Semantic frameworks, MLOps workflows, accessible AI platform. CSIC coordinated.

RAISE/NOA

AI-powered Sun-Earth research infrastructure

First end-to-end European Sun-Earth RI with AI-powered eScience Centre. Event detection, semantic discovery, quality control across heliophysics and space weather.

ACCESS2ACCESS

RI service catalogue with AI-assisted discovery

Recommendations on TNA/VA financial instruments to the EC. Building an RI service catalogue with AI-assisted service discovery. Euro-BioImaging coordinated CSA.

EGI: Towards an AI-Ready Federation

Upcoming AI Services

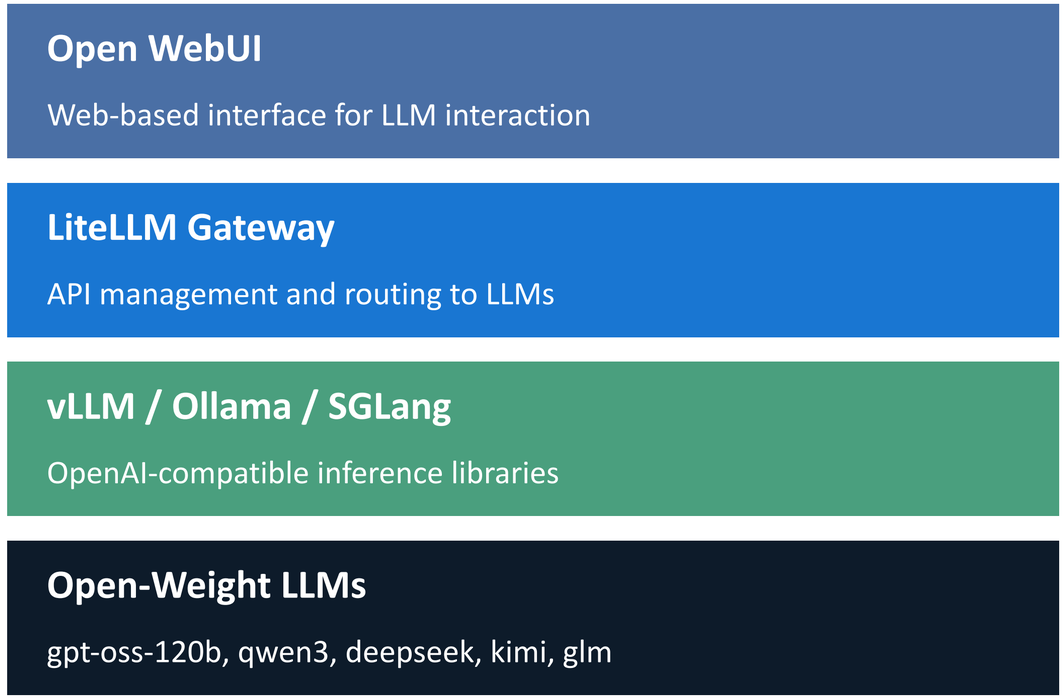
EGI FedLLM – Federated LLM Service

Under incubation | Based on EOSC Data Commons and EOSC Genesis

Use Large Language Models securely with full control over your data

- Secure, high-quality access to open-weight LLMs via web UI and OpenAI-compatible API
- Authenticated through EGI Check-in (AAI)
- RAG workflows and integration into scientific applications
- GDPR-compliant: data not reused for model training
- Two service offers: FedLLM for researchers, FedLLM for communities

Architecture



Current providers: Masaryk University | SZTAKI | CSIC

EGI AI Lab – AI Development Platform

Under incubation | Based on iMagine / AI4EOSC

A platform for developing, training, deploying, and running AI models

- Unified dashboard for compute, storage, datasets, and AI modules
- Full ML pipeline: data preparation to model deployment and monitoring
- Federated learning, development environments, and LLM chatbots
- Provisioned as unified EGI platform or custom instance per community
- Integrated with EGI Check-in, IM, FedCloud, Registry, and Datahub

Who Is It For?

Researchers & Domain Scientists

Develop, deploy, share, and reuse AI modules across disciplines

AI Developers & Projects

Access federated compute and storage for model training and inference

Cloud & HPC Providers

Integrate resources into a federated AI service for the research community

Built on AI4OS (AI for Open Science) | Developed by AI4EOSC consortium led by CSIC

Key Messages

AI is becoming a foundational capability for science, and the infrastructure to support it must evolve accordingly

Open data, open models, and open infrastructure are the foundation for trustworthy AI-driven science

The EGI Federation is evolving from compute federation to AI-ready continuum to serve researchers across borders

Thank You!

Let's Connect

Let’s Connect: www.linkedin.com/in/andreozzi/

For more info: sergio.andreozzi@egi.eu

www.egi.eu

This work is partially funded by the EU research and innovation programme