EUROPEAN OPEN SCIENCE CLOUD: NODE PROPOSAL STATUS

GIACINTO DONVITO

INFN-BARI

AGENDA

- Il contesto / Come siamo arrivati qua

- Che sta succedendo in queste settimane

  - Cosa dice la EC

  - Cosa abbiamo detto noi

  - Cosa hanno detto di noi gli altri

- Roadmap per il futuro

- Considerazioni personali e spunti di riflessione

IL CONTESTO

ꙩꙩeosc Principles for Building the EOSC Federation

The vision for EOSC is to establish a system that enables researchers in Europe to store, share, process, analyse, and reuse FAIR research outputs—such as research data, publications, and software—across disciplines and borders.

- **System of systems approach**: Build on existing and planned infrastructures to create an interconnected ecosystem.
- **Inclusivity and accessibility**: Ensure low barriers to entry, ease of use, and broad participation.
- **Community-driven and participatory**: Engage the research community in co-creating and shaping the Federation.
- **Trust, security, and governance**: Establish clear rules, monitoring, and enforcement mechanisms to ensure a safe and reliable environment.
- **Flexibility and adaptability**: Allow for an evolutionary approach, adapting based on findings and lessons learned.
- **Autonomy and control**: Respect institutional autonomy while ensuring that data and service owners retain full control over shared resources and openness levels.
- **Broad engagement and transparency**: Keep all interested stakeholders informed and involved in key developments.

4

Overview of questionnaire submissions and assessment

Open: 12 Jun 2024 - 18 Aug 2024 (late submissions accepted until 31 Aug 2024)

121 Submissions

25 Countries plus many internationalorganisations

29 Organisations identified to be invited to a dialogue with the Tripartite and EU node.

5 Node types: thematic, national, institutional, regional, and other (such as the EU Node)

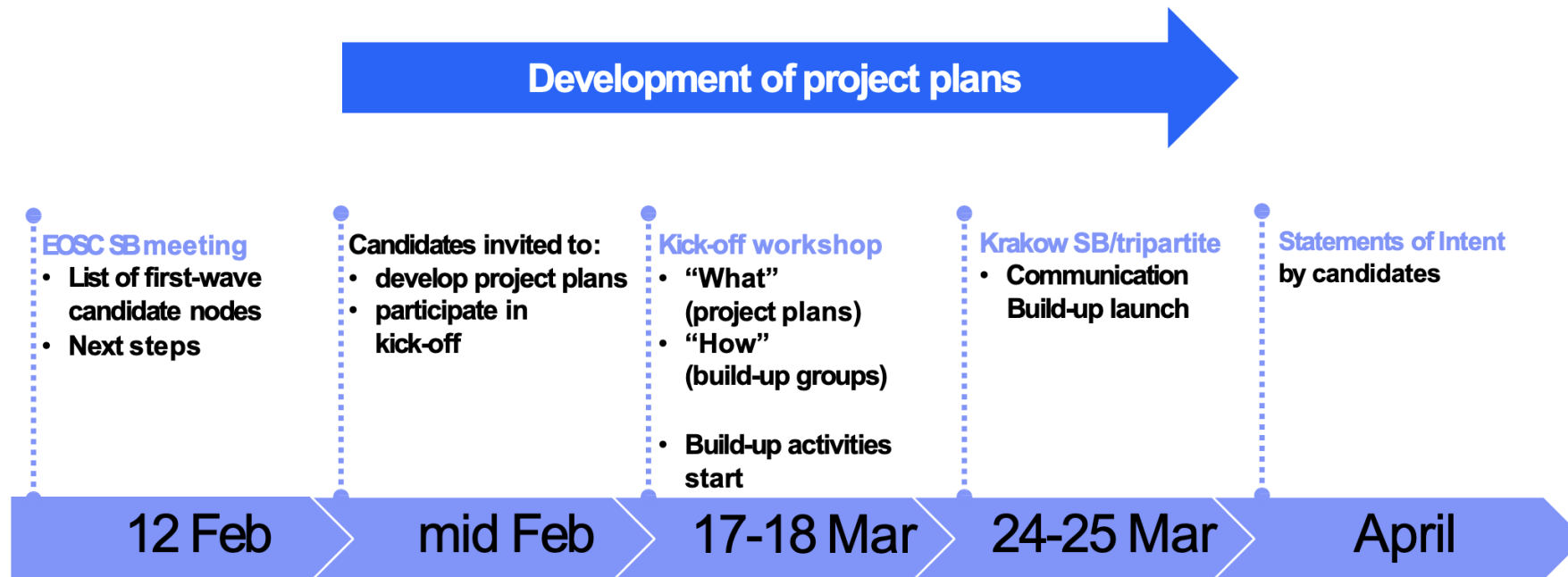**~13** To join the first wave starting March 2025 **+** Subsequent waves of enrolment from Q3 2025 onwards

Goal of dialog process was to get from 29 to ~13 nodes for first wave of deployment

Proposed next steps

Project plans // Kick-off workshop // Build-up Group // Statements of intent

Development of project plans

EOSC SB meeting
- List of first-wave candidate nodes
- Next steps

Candidates invited to:
- develop project plans
- participate in kick-off

Kick-off workshop
- "What" (project plans)
- "How" (build-up groups)

- Build-up activities start

Krakow SB/tripartite
- Communication Build-up launch

Statements of Intent by candidates

| 12 Feb | mid Feb | 17-18 Mar | 24-25 Mar | April |

European Commission

6

13 Candidate Nodes

EOSC Federation Handbook

Purpose of 1ˢᵗ version of Handbook

1 How to build the EOSC Federation during the period 2025 to 2027

2 What will the EOSC Federation enable and provide for researchers

3 How will the success of the EOSC Federation be measured
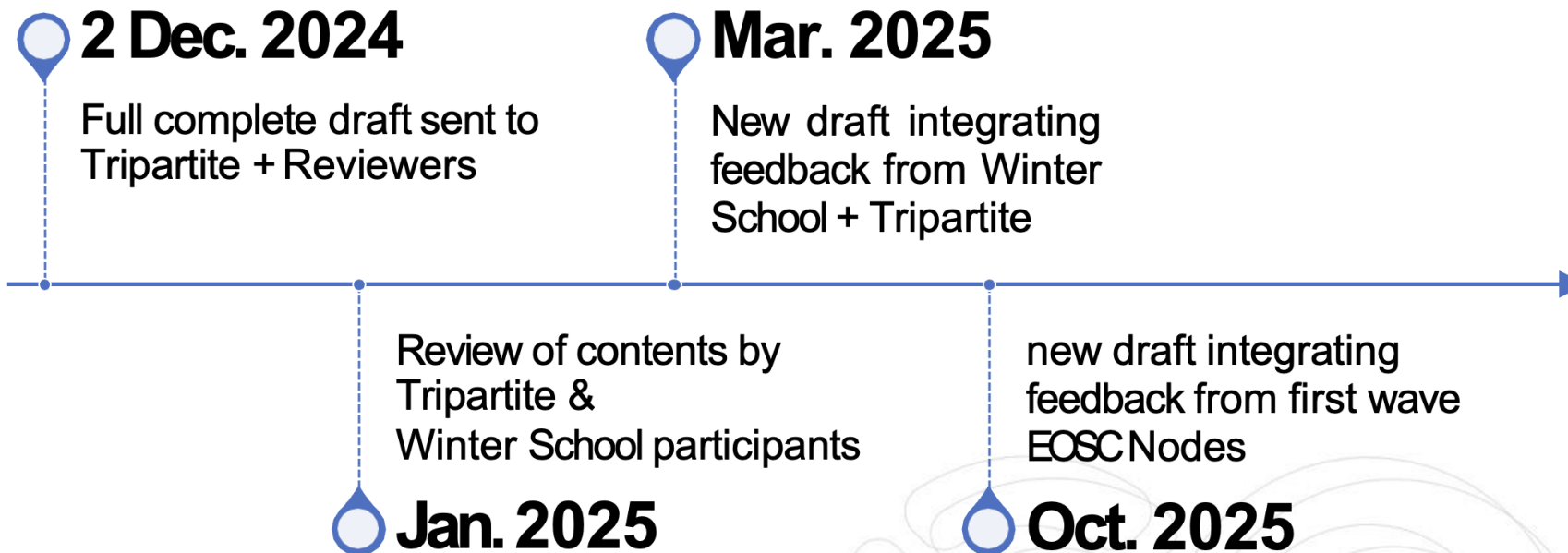
Outline of EOSC Federation Handbook (60+ pages)

- Chapter 1 – Purpose
- Chapter 2 – Governance
- Chapter 3 – Operations

- Chapter 4 – Architecture
- Chapter 5 – Scientific Resources
- Chapter 6 – Policies

EOSC Federation Handbook status

First full draft now available

2 Dec. 2024

Full complete draft sent to Tripartite + Reviewers

Mar. 2025

New draft integrating feedback from Winter School + Tripartite

Review of contents by Tripartite & Winter School participants

Jan. 2025

new draft integrating feedback from first wave EOSC Nodes

Oct. 2025

STATO DELLE ATTIVITÀ

IL PUNTO DI VISTA DELLA EC

Objectives of the workshop

The kick-off workshop shall serve to commonly...
- ...get to **know each other**, find gaps and commonalities.
- ...elaborate the planning of the **main activities of the build-up phase**, on the basis of **draft project charters** prepared by the organisations and on the input by the **EU Node representatives**. This includes **prioritization of use cases and tasks, developing a common timeline.**
- ...set out the **organisational modalities of the build-up phase** and define the **next steps**.

EOSC Federation build-up phase aims for Q4 2025

EOSC Federation will:
- Be comprised of several enrolled EOSC Nodes.
- Have operational governance and a clear organizational structure.
- Have established key policies and the EOSC Interoperability Framework.
- Demonstrate cross-Node utilization of federating capabilities.

EOSC Federation users will be able to:
- Access resources via single sign-on and a common catalogue.
- Use a virtual research environment.
- Access FAIR and machine actionable research datasets .

Communities/stakeholders will be provided with:
- Key information via the EOSC Federation Handbook.
- Guidance on Node enrolment and resource onboarding.
- Demonstrations of multi-node use cases.

Workshop 17-18 Marzo Bruxelles

3.2.3 Roles for EOSC Nodes

In order to facilitate the interaction of the Node with the Federation, each Node must nominate a number of persons in the following roles:

1. A **Coordinator** who represents the Node in the Federation.
2. An **Operation Manager** who is the contact person with the Operations Team of the Federation on operations.
3. A **Technical Coordinator** who is the contact person with the Technical Coordination Team of the Federation on technical issues (mainly IT issues).
4. A **Security Officer** who is the contact person for all security related issues.
5. Thematic Nodes will nominate a **Scientific Officer** who represents the Node's scientific content.
6. A **Legal/Privacy Officer** for legal issues

COSA ABBIAMO DETTO NOI

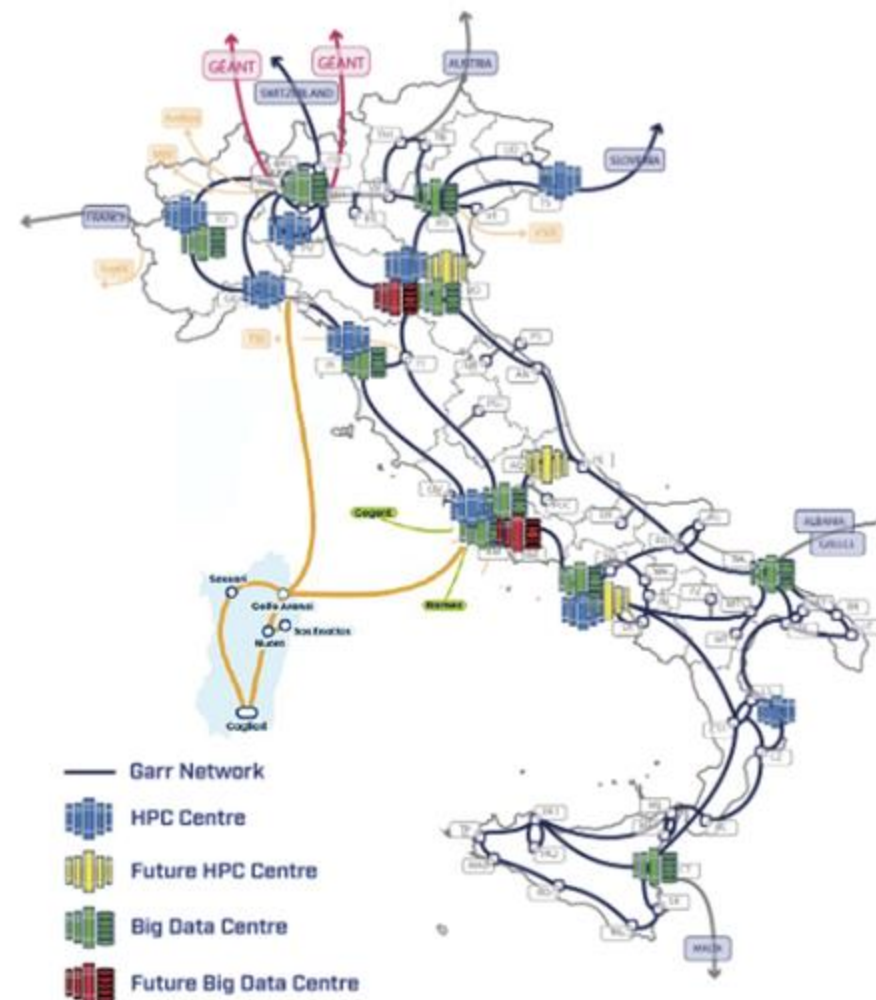# EOSC Italian Node – ICSC
Facts & figures

The **EOSC Italian Node** (IN) is managed by the **ICSC Foundation**, a legal entity with defined governance, stable organization, resources and expertise to guarantee the deployment and integration of IN activities/resources by Q2 2025.
ICSC federates resources of 50+ members including leading RPOs, e-Infra and RI
The application as the Italian EOSC Node has been in line with the active participation of the Italian Computing and Data Infrastructure (**ICDI**), which was one of the four **founding members of EOSC** and has coordinated Italy's participation in EOSC since 2018.

Founded in 2022, ICSC operates the EOSC Italian Node as a federated infrastructure providing services that have been active for 20+ years. ICSC/IN is committed to contributing to the EOSC Federation beyond 2025/26 and is strongly **supported by the Italian government** in **alignment with the European strategy**.

ICSC manages a **shared, federated and open infrastructure**, representing a unique strategic asset for Italy and for the international community, supporting Open Science and EOSC.

— Garr Network
HPC Centre
Future HPC Centre
Big Data Centre
Future Big Data Centre

EOSC Italian Node – ICSC
Organizational capacities

- The EOSC Italian Node (IN) will provide 4 FTEs for 2025, with the following roles:
    - Coordinator: Matteo Zanaroli
    - Technical Officers: Paolo Manghi and Giacinto Donvito
    - Operation Officer: Alessandro Costantini
    - Scientific Officers: Mauro Campanella and Davide Salomoni

- In the build-up phase, additional roles and responsibilities will be defined and additional FTEs will be made available to cover i.e. security, user onboarding, training, communication activities…

- EOSC IN will also benefit of the long-term active contribution of 29 Italian esperts participating in the EOSC Association Task Forces and in EOSC projects.

EOSC Italian Node – ICSC
Core Services and Federating capabilities

Functions	Status	Description
An **integrated front-end portal**	**BETA released (TRL8, to go public in March 2025)**	• Single entry-point to EOSC Italian Node Core services and National services and resources
Service Registry/Catalogue	**BETA released (TRL8, to go public in March 2025)** • Compatible with EOSC EU Node schema • To be integrated with EOSC EU Node	• Enables registration of services by National service providers • The process is curated and mediated by administrators • Onboarding policies to be aligned with the EOSC EU Node catalogue
Research Product Registry/Catalogue	**Delivery expected soon** • To be federated with OpenAIRE Graph and visible via EOSC ResourceHub • Richer curated metadata, full coverage of research products • Supporting EOSC policies to Italian data sources	• Federating registries/catalogues and data sources operated by Italian institutions, including Universities, Research Centers, and RIs • Many Italian data sources are registered in the OpenAIRE Graph and visible via the EOSC ResourceHub
AAI	**Operational (TRL9)** • Distributed throughout the ICSC Data Lake, generally adopted by National services to enable access	• Integration with eduGAIN, IDEM, LifeScience-AAI, EOSC.eu node and others. • Consistent authorization management via OAuth Tokens. • Integration of AuthN policies via Open-ID Connect, SAML • Ongoing integration with EuroHPC Centers (CINECA).
Service Monitoring and Accounting	**Operational (TRL9)**	• Tracking availability of services in the catalogue • Tracking access to services in the catalogue by authorized users
National Helpdesk	**Delivery expected soon** • Reusing EOSC Beyond helpdesk software packages so as to be federated by the EOSC EU Node Helpdesk	• Support EOSC Italian Node users for EOSC Italian Node services and Helpdesk aaS for Italian Service providers
SMS	**Delivery expected soon** • The solution will adopt an ITIL compatible standard	• CSC/IN offers access to services, systems, data and infrastructure from several RPO and SPO operating under a variety of SMSs ITIL compatible standards

All the Italian Node Core Services will be operated by ICSC members according to SLAs and ToUs between service providers and ICSC

EOSC Italian Node – ICSC

Enabling federating capabilities for computing and storage for science at National level

- A **full Compute and Data Management stack** with:
  - Federated **distribution of compute services** via a PaaS-level orchestrator, connecting multiple cloud sites.
  - A **modular portfolio of compute services**, defined through standards-based TOSCA templates (to be tested with the EOSC EU Node).
  - **Data management policies** for automated data replication, reliable file transfer of multi-tenant data (with proven scalability to at least hundreds of PB data in the case of large physics experiments but applicable to any scientific domain).

- **Distributed cloud and data processing capabilities** based on OpenStack, Kubernetes (compute) and multiple storage technologies (Posix, S3, Webdav, etc).
  - Capability of leveraging HPC resources, with transparent access connected with Cloud services

- **Support for providing analytics, storage and transfer of sensitive data through ISO-27001, 27017, 27018 certified Cloud regions** to support use cases involving data confidentiality, integrity, availability, with GPDR compliance.

EOSC Italian Node – ICSC contribution to federating capabilities

Contribution to the EOSC Federation in the build-up phase

ICSC/IN National cloud supercomputing and storage federated infrastructure will make available to EOSC Federation users (under specific SLA):

- 25 PFLOPs HPC

- 8k cloud/grid cores

- 11 PB tape storage, 15 PB disk storage

Data Lake and HPC/Cloud federation of data centers

- ISO27001 certified cloud regions

Data transfer services benefits of GARR (Italian NREN) connection with hundreds of Gigabits per second.

EOSC Italian Node – ICSC contribution to federating capabilities
Towards an EOSC AI ecosystem

- **Federating resources in support of AI experiments** through direct participation to Exascale Italian AI Factory (ICSC leads the Data Services and Training / Capacity Building WPs)

- EOSC IN will offer **ELSI (Ethical, Legal and Social Implications) services to the Federation under specific agreements**, including the implementation of Sandboxes for AI
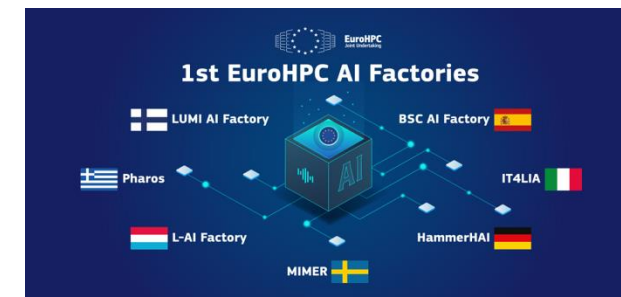
EOSC Italian Node – ICSC
Multi-node use cases - CERN Indico

- **Main goal**: Deploy an **Italian instance of CERN's INDICO Agenda Service**, exploring the realization of a federation of INDICO services through common AAI solutions.

- **Value proposition**:
  - Provide the Italian community with an Indico Agenda Service and create a framework for other nodes towards a federation of INDICO services.
  - Create a potential backup and disaster recovery solution for the EU-Node Indico Services.

- **In-scope**: Deploy an Italian instance of CERN's INDICO Agenda Service and a proposed architecture and solutions to provide backup and disaster recovery for the EU-Node services. Providing best practises for implementing national (or thematic) instance of Indico service.

- **Out of scope**: Further development of the Indico software.

EOSC Italian Node – ICSC
Multi-node use cases - CERN Reana

- **Main goal**: Exploit CERN's REANA (Reusable and reproducible research data analysis platform) to **instantiate computational workflows on heterogeneous remote clouds**.

- **Value proposition**:
  - Provide to the Italian community towards a European federation of experimental workflow environments decoupled from cloud and resource providers

- **In-scope**: Deploy an Italian instance of REANA, definition of policies and customization for Italian Research Community. Share experience and know-how with other nodes.

- **Out of scope**: Further development of the REANA software.

EOSC Italian Node – ICSC
## *Multi-node use cases - BBMRI*

- **Main goal**: **Deploy a cloud-based use of interactive notebooks**, with federated data access involving also integrated use of HPC-enabled distributed resources to support AI based algorithms.

- **Value proposition**:
  - Exploit sensitive data storage and analysis facilities using the ICSC ISO-certified cloud regions.
  - Provide compute and AI capacities in the European Health Data Space (EHDS) compliant manner for data-users to the research community leveraging ICSC resources.

- **In-scope**: Exploit sensitive data storage and analysis facilities and provide processing capacities for data-users in EHDS Secure Processing Environment (SPE) compliant manner.

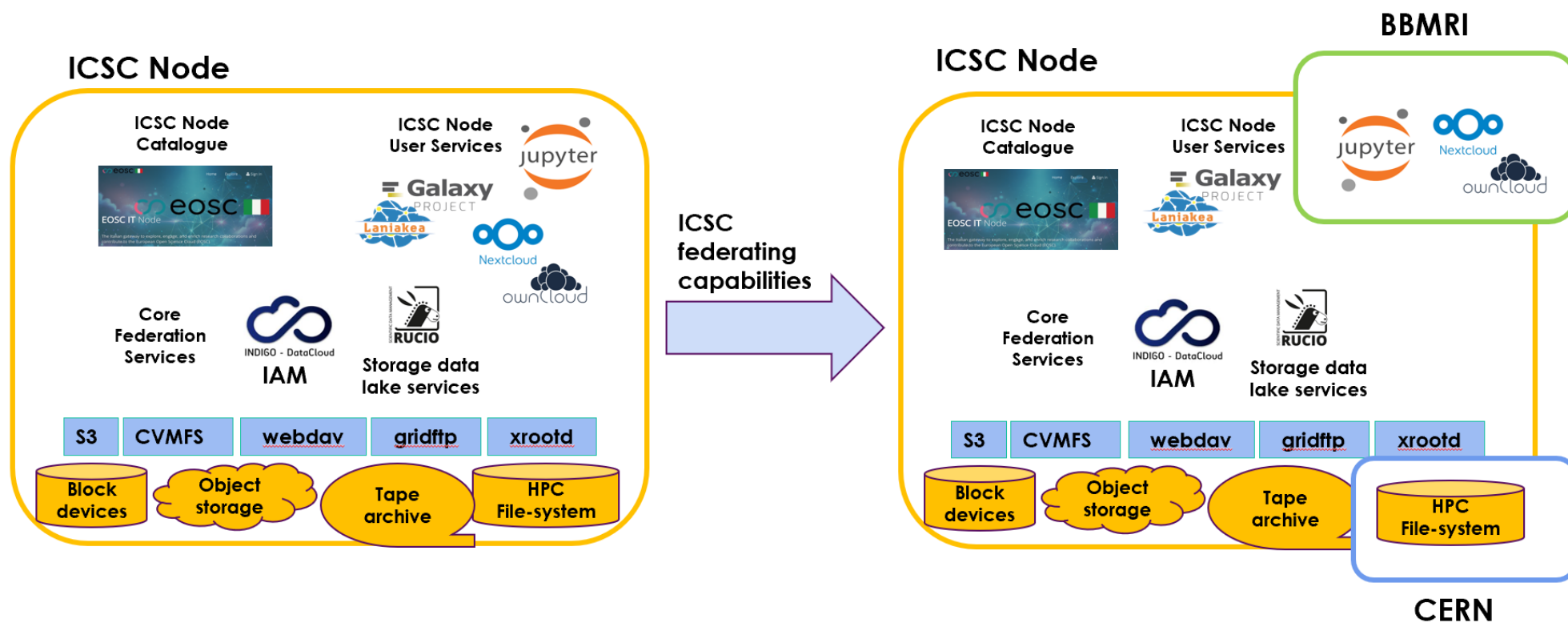- **Out of scope**: Further development of BBMRI services.

EOSC Italian Node – ICSC
Federation of computational resources and services
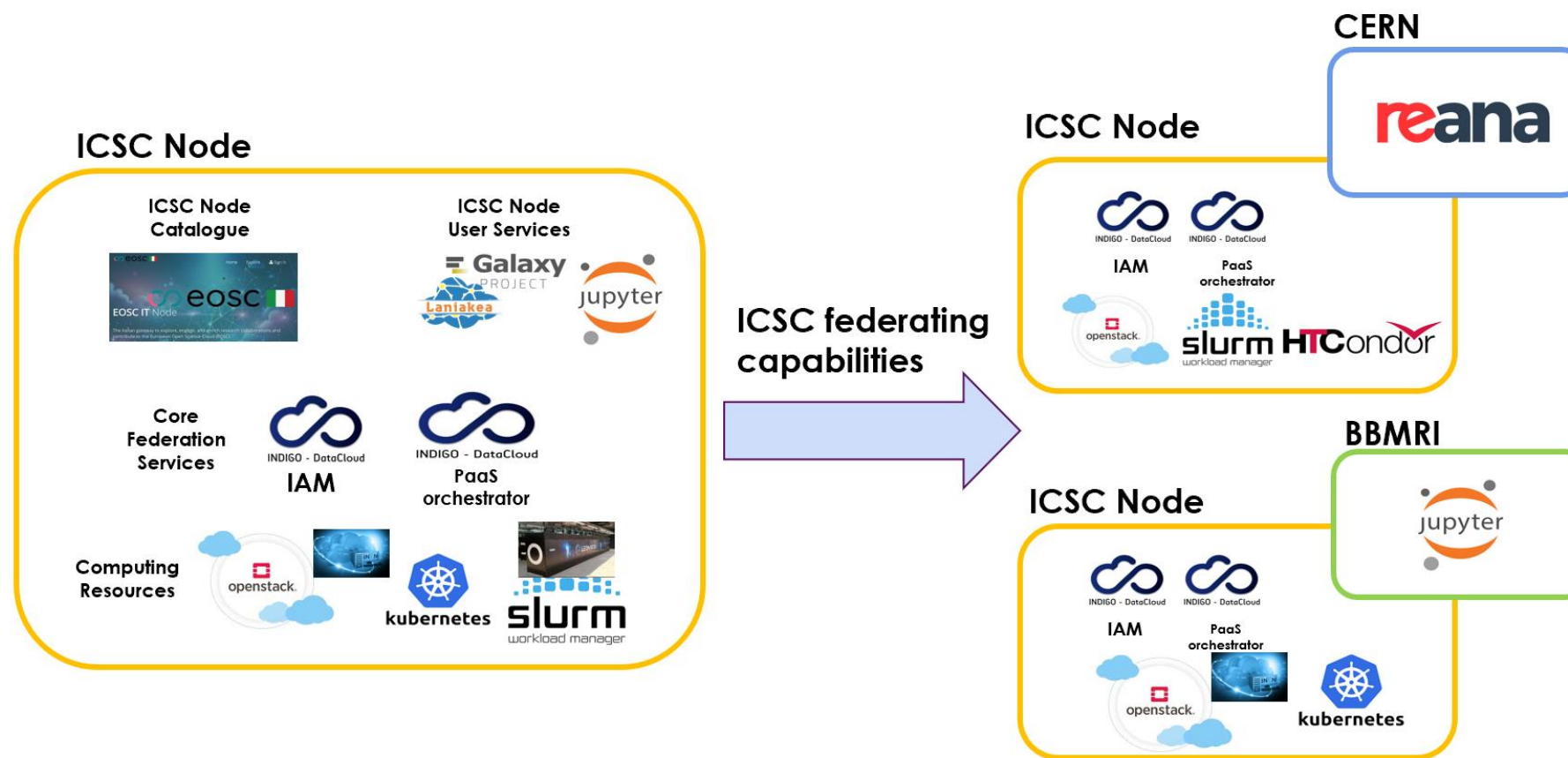
EOSC Italian Node – ICSC
*Federation of **data** resources and services*

EOSC Italian Node – ICSC
External dependencies and risks

General dependencies and risks

- AAI at Federation level

- Interoperability and technical specification with EOSC EU-Node

- Sustainable collaboration between the shareholders of the Federation

- Consensus on adopted policies

- Underestimation of manpower effort

Use-case/domain specific dependencies and risks
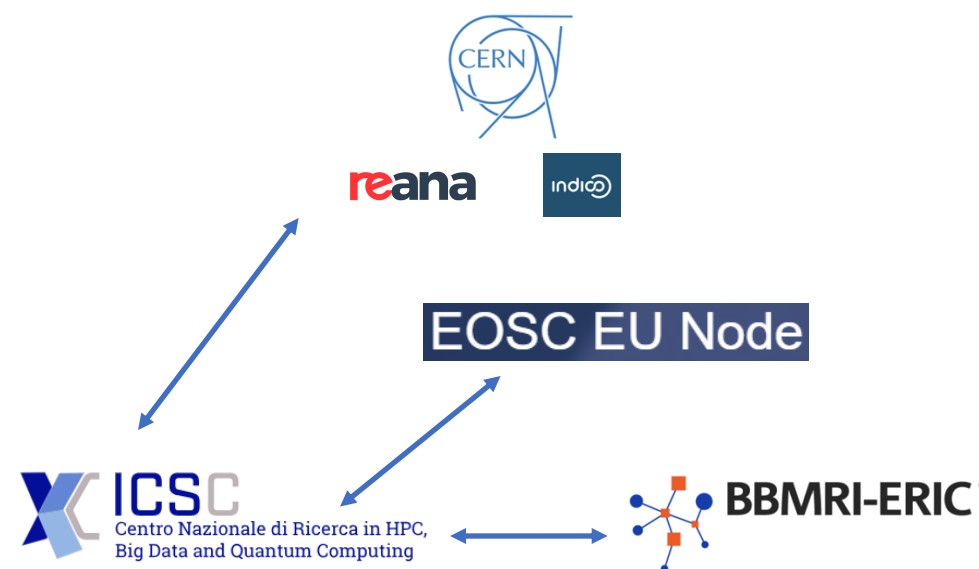
EOSC Italian Node – ICSC
Added value

- Contribute to change the culture and **make Open Science the new normal**
- Support a **strong federation** of infrastructures for data-driven research
- Contribution to the **EOSC developments and policies**
- Definition of a credit system for a federation of cloud-based interoperable services
- Participation in **projects aimed at developing services to exploit the full federation of EOSC Nodes**
- Increase the **opportunities for Italian researchers** and research communities
- Having and giving access to high value multidisciplinary competencies and datasets, increasing the quality of heterogeneous data for AI
- Develop education and training programs
- **Synergy** with EuroHPC key initiatives such as the **AI Factory**, and enhanced federation with distributed **HPC, Quantum & Cloud facilities**

# EOSC Italian Node – ICSC
Timing and milestones

Milestone	Expected Month	M1 Apr '25	M2 May '25	M3 Jun '25	M4 Jul '25	M5 Aug '25	M6 Sep '25	M7 Oct '25	M8	M9	M10	M11	M12
Project Charter	May 2025												
Analysis of the EOSC EU-Node federating capabilities and Plan for integration	July 2025												
EOSC Italian Node Governance and Management	June 2025												
Stakeholder and community engagement strategy	June 2025												
Interoperability policies and procedures - first version	September 2025												
EOSC Italian Node core services - first wave	September 2025												
EOSC Italian Node core services - second wave	TBD												
Use-cases - first results	October 2025												
Use-cases - further developments	TBD												

COSA HANNO DETTO GLI ALTRI DI NOI

Use case 2: Multi-centric validation of AI model for prostate-cancer screening

- ▶ **Goals:** Prostate screening model is developed using data only from BBMRI.cz/MMCI and is validated with data from BBMRI.at/MUG
- ▶ **Added value:**
  - ■ Enable multi-centric validation of AI models for medical researchers
  - ■ Enable multi-centric retraining of AI models
- ▶ **Dependencies:** Masaryk Memorial Cancer Institute, RationAI at Masaryk University, Medical University Graz, EOSC EU Node, ICSC, possibly other HPC providers

Cancer probability 99,5%

This is where the cancer is found by the model
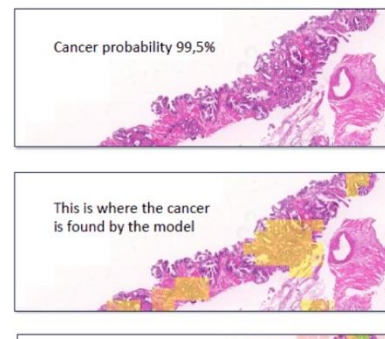
Use case 3: AI-Guided Biomarker Discovery for CRC Risk Stratification

- ▶ **Goals:**
  - ■ Develop digital biomarker for colorectal cancer stage 2 stratification
  - ■ Demonstrate entire governance for compliance with the CEN/CENELEC AI standards in the Life-Science and HPC domains
- ▶ **Added value:** Enable computation compliant with the CEN/CENELEC AI standards for medical researchers
- ▶ **Dependencies:** Masaryk Memorial Cancer Institute, RationAI at Masaryk University, Medical University Graz, EOSC EU Node, ICSC, possibly other HPC providers

Workshop 17-18 Marzo Bruxelles

CERN Node - A two-seasons story

1: EU Node Integration

CERN's participation in EOSC strengthens the EU Node with high-TRL services.

Learning experience for Phase 2

Deploy/Integrate CERN Services:

2: CERN node

→ Will investigate the EOSC Node 'builder kit' announced by DG CNECT as a basis for dashboard
→ Offer limited resources for other domains to exploit HEP Data

# Use case 1 - reana Reproducible analysis workflows

Several sub-use cases:

2nd) Involves **generic EOSC Node interested in deploying REANA**

→ Enables execution of containerized workflows close to the data

→ Can be integrated into EOSC services such as Zenodo, and ResourceHub

→ in discussions with ICSC, PUNCH/NFDI (future?)

→ would open door for future hybrid distributed workloads
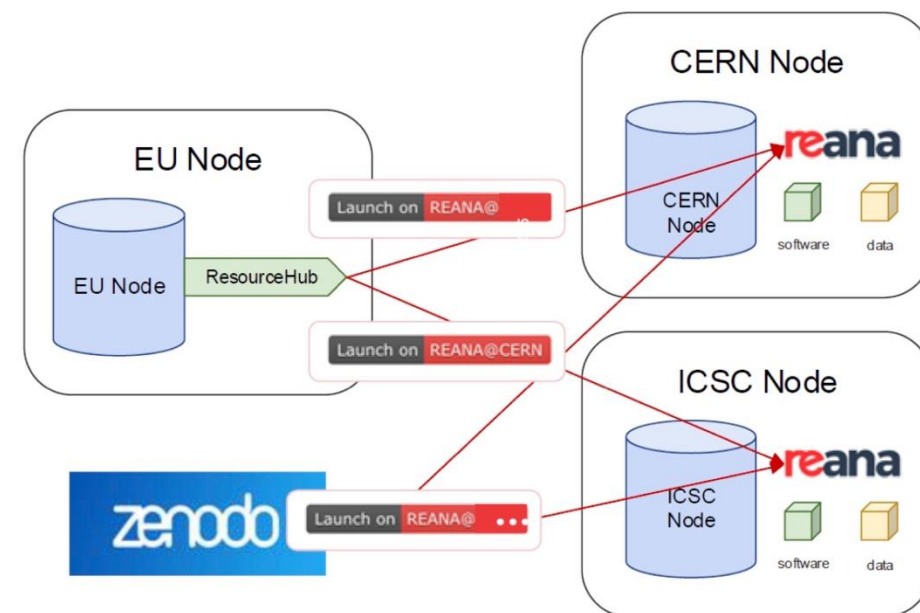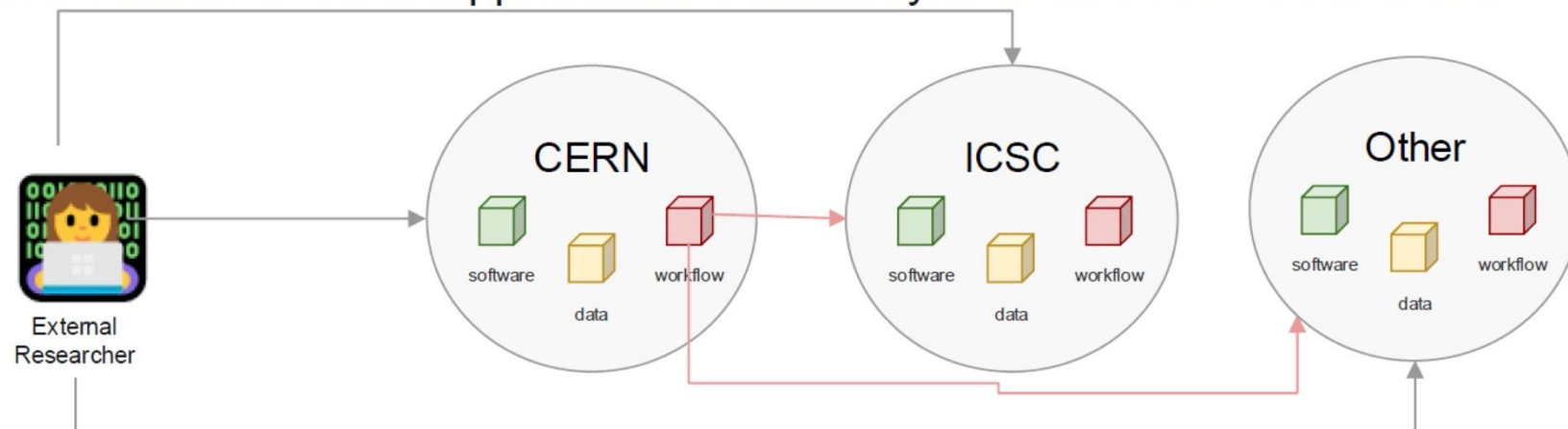
Added Value for End Users

Extended access to particle physics open data and software

Interdisciplinary collaboration: scientists from diverse fields can benefit

Efficient use of resources: reduces need for large data transfers

Enables scalable, reproducible research: **workflows run on trusted infrastructure**
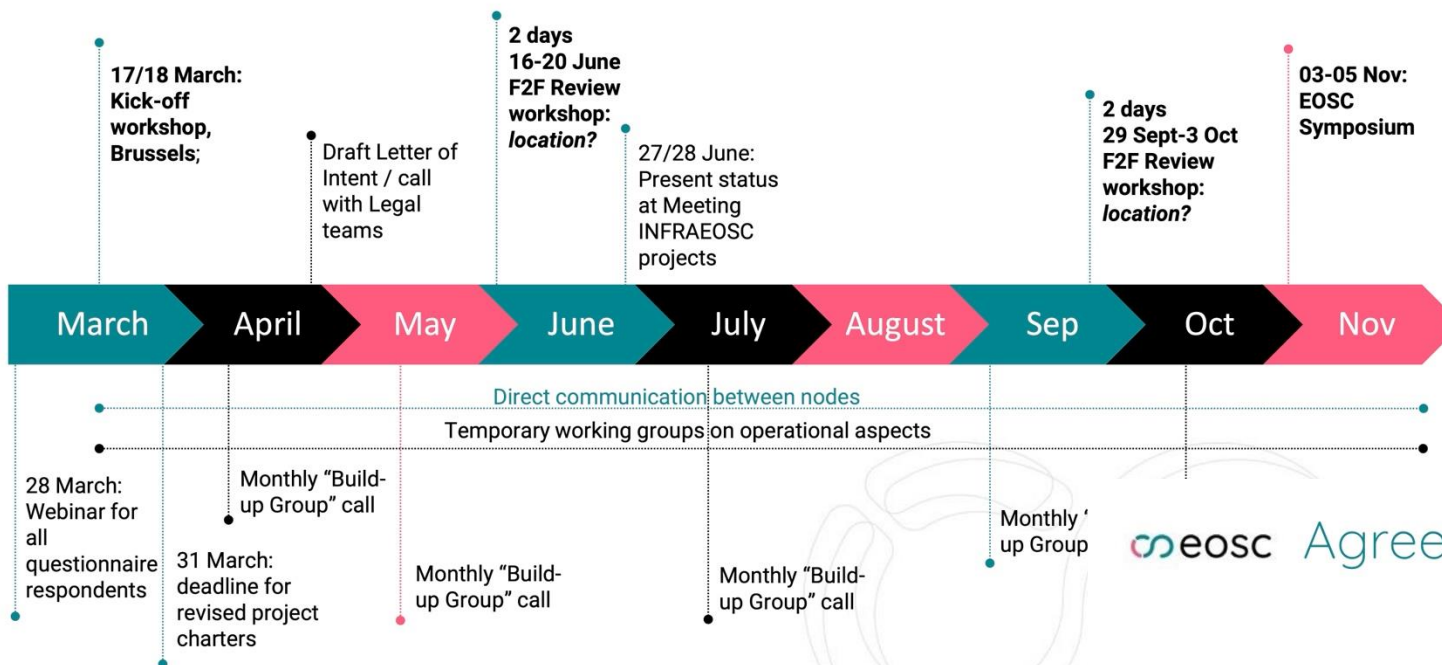
Federated workflows: could support multi-node analysis across EOSC federation

ROADMAP PER IL FUTURO

Indicative Timeline for 2025

17/18 March: Kick-off workshop, Brussels;

Draft Letter of Intent / call with Legal teams

2 days 16-20 June F2F Review workshop: *location?*

27/28 June: Present status at Meeting INFRAEOSC projects

2 days 29 Sept-3 Oct F2F Review workshop: *location?*

03-05 Nov: EOSC Symposium

March	April	May	June	July	August	Sep	Oct	Nov

Direct communication between nodes

Temporary working groups on operational aspects

28 March: Webinar for all questionnaire respondents

31 March: deadline for revised project charters

Monthly "Build-up Group" call

Monthly "Build-up Group" call

Monthly "Build-up Group" call

Monthly "Build-up Group" call

Monthly "up Group"

Agreeing next milestones

31.03.	Submission on revised / update Project Charters based on discussion today's discussions
31.03.	Scoping of Sub-Groups
01.-04.04.	First Build-up Group meeting / identification of Co-Chair
04.04.	Agreement of Sub-groups
Week 16.06.	F2F workshop
Week 29.09.	F2F workshop

18 March 2025

eosc

Symposium
2025

03–05 November
Brussels

CONSIDERAZIONI PERSONALI E SPUNTI DI RIFLESSIONE

- Sembra evidente che le evoluzioni successive dei finanziamenti e degli sviluppi saranno correlati alla creazione/federazione di Nodi EOSC.

  - Vedi call EOSC-01 nella presentazione di Luciano Gaido

- Attualmente tutto l'effort è «in-kind»

  - Ma appunto la prima fonte di finanziamento certo «deve» essere la EOSC-01

  - Chi sta nella prima tornata di nodi, detta le regole per il futuro

- Tecnicamente noi siamo fra i nodi più «forti»

  - Nella valutazione della seconda fase siamo fra i pochi nodi che hanno preso solo «stelle verdi»

  - Anche il fatto che abbiamo già una federazione (ICSC) che rappresenta il calcolo scientifico nazionale, è una singolarità da sfruttare in ambito EU

CONSIDERAZIONI PERSONALI E SPUNTI DI RIFLESSIONE

- Dobbiamo decidere / capire le relazioni con ICSC e gli altri attori Italiani (CNR, CINECA, etc)

- L'iniziativa EOSC – Node non è isolata: è strettamente collegata con le altre attività sul calcolo di ICSC/INFN/CINECA, etc:

  - ➔ La AI factory:

    - Cosa vogliamo / possiamo fare come INFN in questo contesto?

  - ➔ I nuovi «PON»:

    - Sono appena uscite 2 nuove call, come partecipiamo? Per fare cosa?

  - ➔ I finanziamenti su HPC

COME SARÀ IL FUTURO??

Ecco un'immagine generata con DALL·E che rappresenta una visione futuristica della European Open Science Cloud (EOSC).

Mostra un cloud digitale avanzato che collega istituzioni di ricerca europee attraverso flussi di dati olografici, simboleggiando collaborazione, innovazione e accesso aperto alla conoscenza.