Launching and operating the EOSC EU Node

approach, milestones, and the role of the broader community

DG CNECT Unit C.1 Open Science and Digital Modelling

Peter Szegedi

What is EOSC

A process

- Accelerate Open Science, FAIR data management and use of digital methods and services
- Stimulate co-operation in science and research, new insights and innovations, higher research productivity and improved reproducibility in science.

An open, trusted, federation of infrastructure

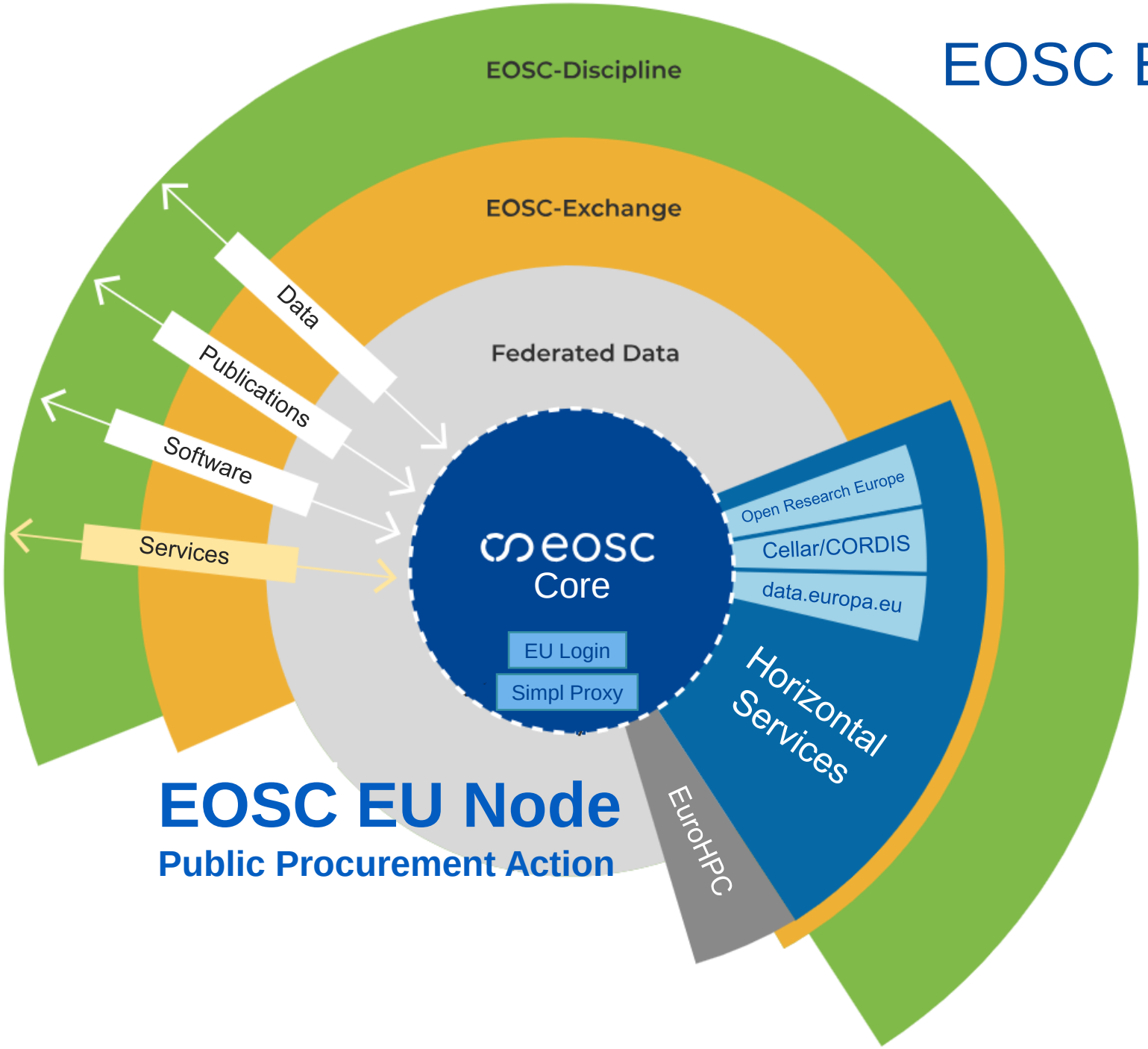
- Access existing Research Infrastructures in Europe;
- Enable circa 2 million European researchers to store, share, process, analyze, and reuse research digital objects (e.g. data, publications and software)

An evolving ecosystem

- Bringing together the European Commission, the governments and the many R&I stakeholders involved in the European Research Area
- Co-created across European, national, and institutional levels

European Commission

EOSC EU Node Value Proposition



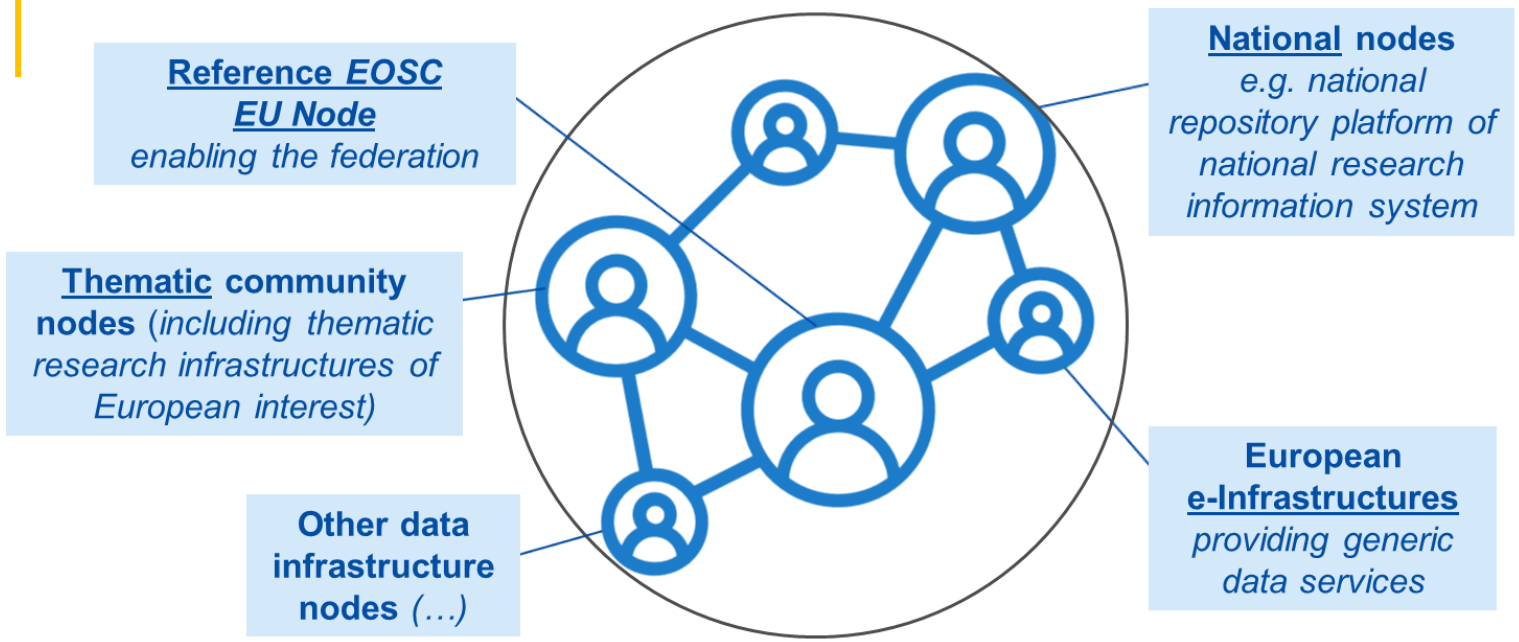
EOSC EU Node

Public Procurement Action

- **Facilitate** the creation of the *"Web of FAIR data and interoperable services"* (aka. EOSC Federation) under the Open Science Policy
- **Put** a *"seed in the ground"* by operationalizing the first recognised EOSC Node at the European level for the initial 3 years
- **Offer** "*core services"* for scientific research infrastructures to federate (single-sign-on, catalogues, knowledge graph, application workflow, monitoring, accounting, helpdesk) and common "*horizontal services"* for end-users to benefit from (compute, containers, data transfer, notebooks, file sharing, open research data)
- **Define** the *pathway and blueprint* (EOSC Interoperability Framework) for other potential EOSC Node operators to join the federation

European Commission

About the EOSC Federation and EOSC Nodes



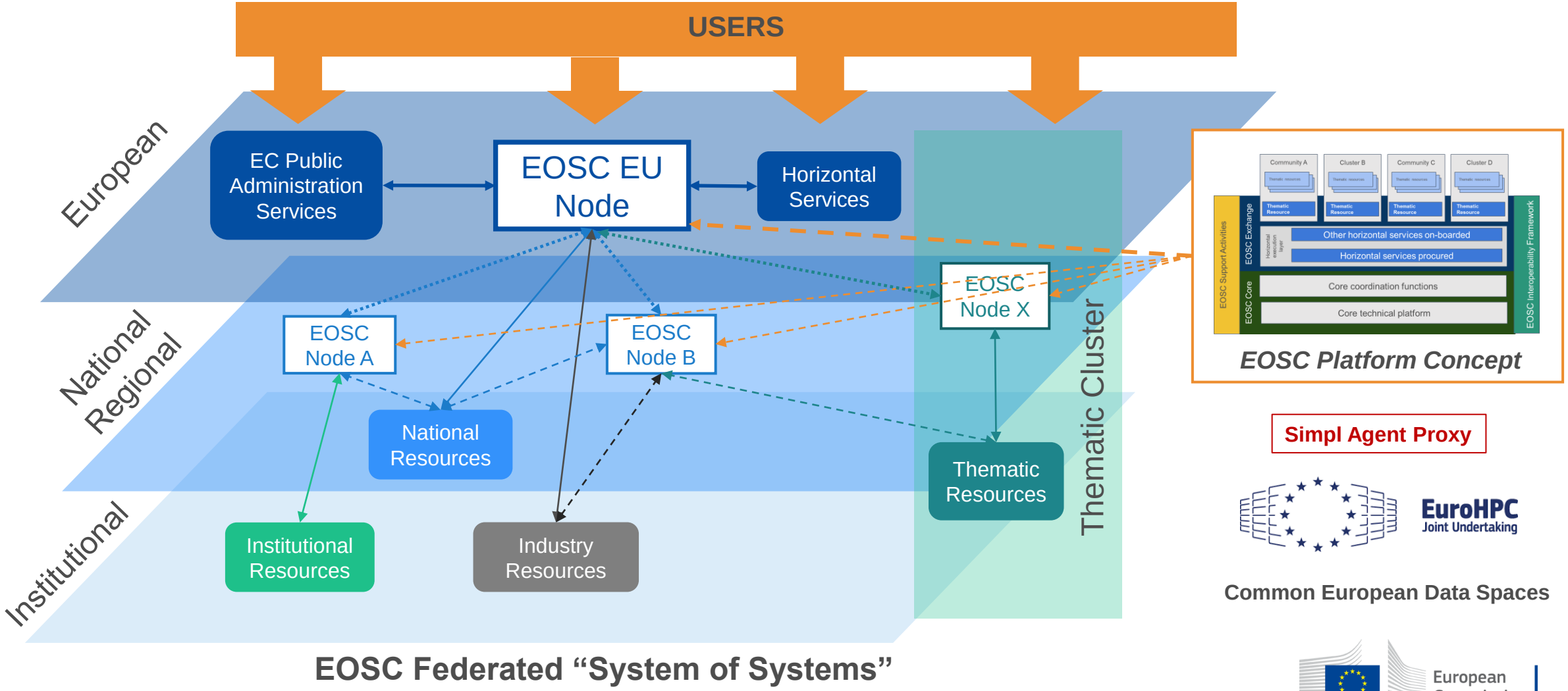
EOSC policies and standards**: A baseline should be defined to ensure that each node can have a minimum working set of features and supports a minimum set of policies. It is important to mandate compliance with protocols and standards, but to give freedom to each node on how to support them.

EOSC Federation*: Open and trusted federation of collaborative, autonomous infrastructures applying agreed, consensus-based EOSC policies and rules of participation, combined into a system of systems to enable European researchers to store, share, process, analyse, and reuse research digital objects (e.g. data, publications and software)

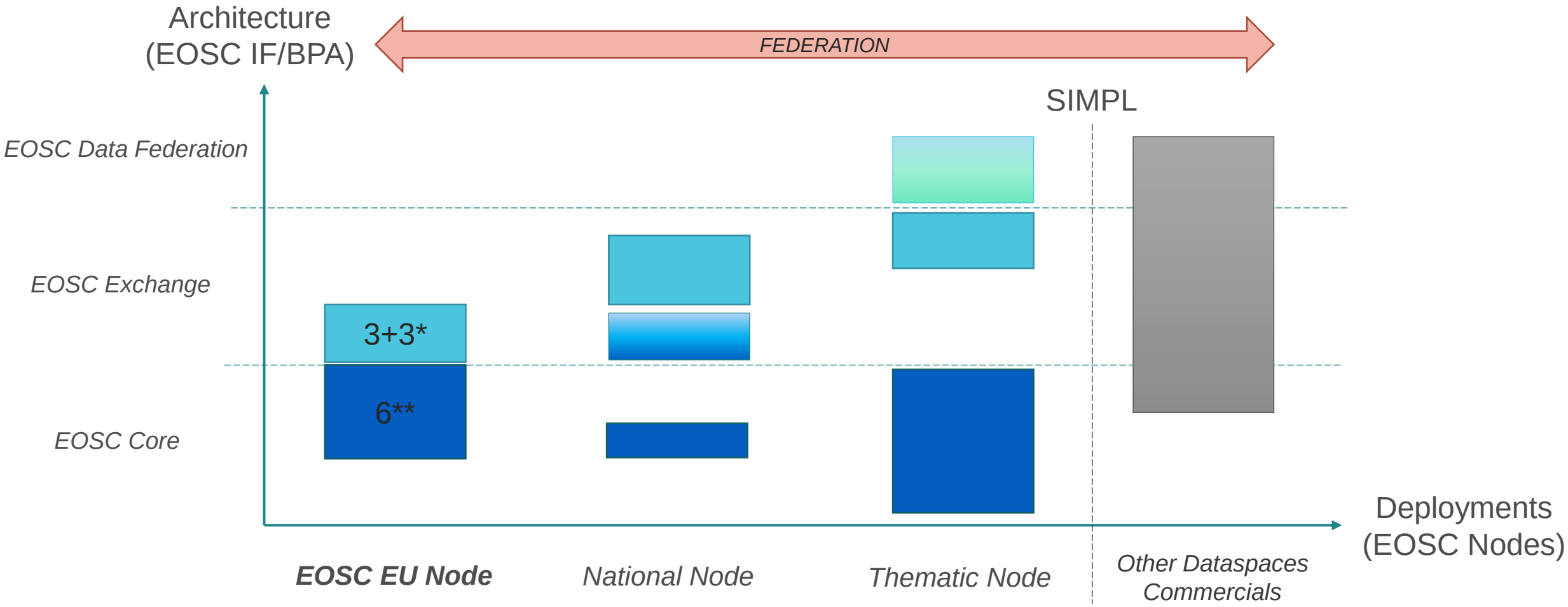
EOSC Node*: Data infrastructure system of variable nature (national, regional, institutional or thematic) with consensus-based policies, transparent ownership and clear responsibility, connected to the EOSC Federation to share information and resources within the EOSC community and to leverage common services

* Source: "*EOSC operations and evolution post-2027*" supporting document by the EOSC-SB Policy subgroup (November 2023)
** Source: GEANT and NREN's position on EOSC Nodes (October 2023)

EOSC EU Node – Federation Approach



Architecture vs. Deployment models



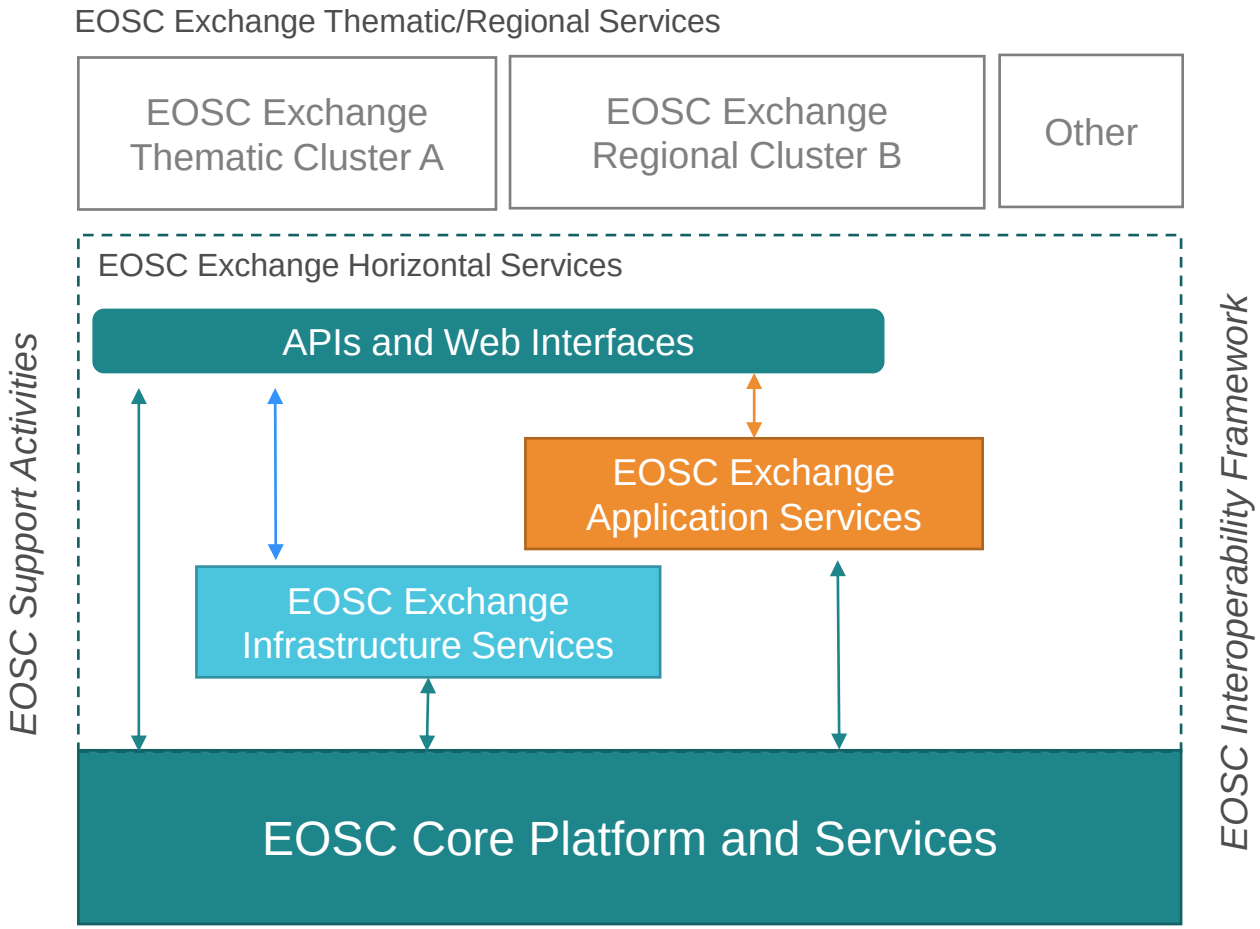
*3+3 **EOSC Exchange Horizontal Services:** Virtual Machines, Containers, Bulk Data Transfer, File Sync&Share, Notebooks, Large File Transfer
6 **EOSC Core Services: Authentication and Authorization/SSO, Application Workflows, Resource Catalogues, Monitoring, Accounting, Service Management

European Commission

EOSC EU Node characteristics and features

- **European level multi-disciplinary and multi-national** scientific data/service portfolio for all research users (eduGAIN) and citizen scientists (EU Login/eIDAS)
- **For now, owned by the EC** and governed by the EOSC Tripartite Governance (EC, EOSC-A, MS/AC) Future ownership is under discussion.
- **Operated and maintained 24/7** by contracted third-parties (result of the EOSC Procurement Action) in production
- **SIMPL Agent proxy** to connect to other industrial Data Spaces
- **EuroHPC resources** may be offered to the EOSC Federation
- **Open concept: National, regional and/or thematic** service providers as well as autonomous EOSC Nodes can connect to the federation (established interoperability frameworks and policies)

Hight-level EOSC EU Node architecture



Procurement Lot Structure

Lot 3	Managed Collaborative Data Platform, Interactive Data Analytics Platform and Visualization Services for the EOSC Exchange (Application Services)
Lot 2	Managed Container Platform and Virtual Machine Services for the EOSC Exchange (Infrastructure Services)
Lot 1	Managed Services for the Development, Integration, Deployment and Operations of the Federated EOSC Core

PSNC

Research & Innovation Information Technologies / ATHENA

European Commission

All together the awardees

EOSC EU Node

EXCHANGE

PSNC

safespring

NORDUnet
Nordic Gateway for Research & Education

Sikt

egi
cesnet

GÉANT

ownCloud

CORE

Research & Innovation
Information Technologies
ATHENA

OpenAIRE

GWDG
Gesellschaft für wissenschaftliche Datenverarbeitung mbH Göttingen

egi
cesnet

GÉANT

grnet

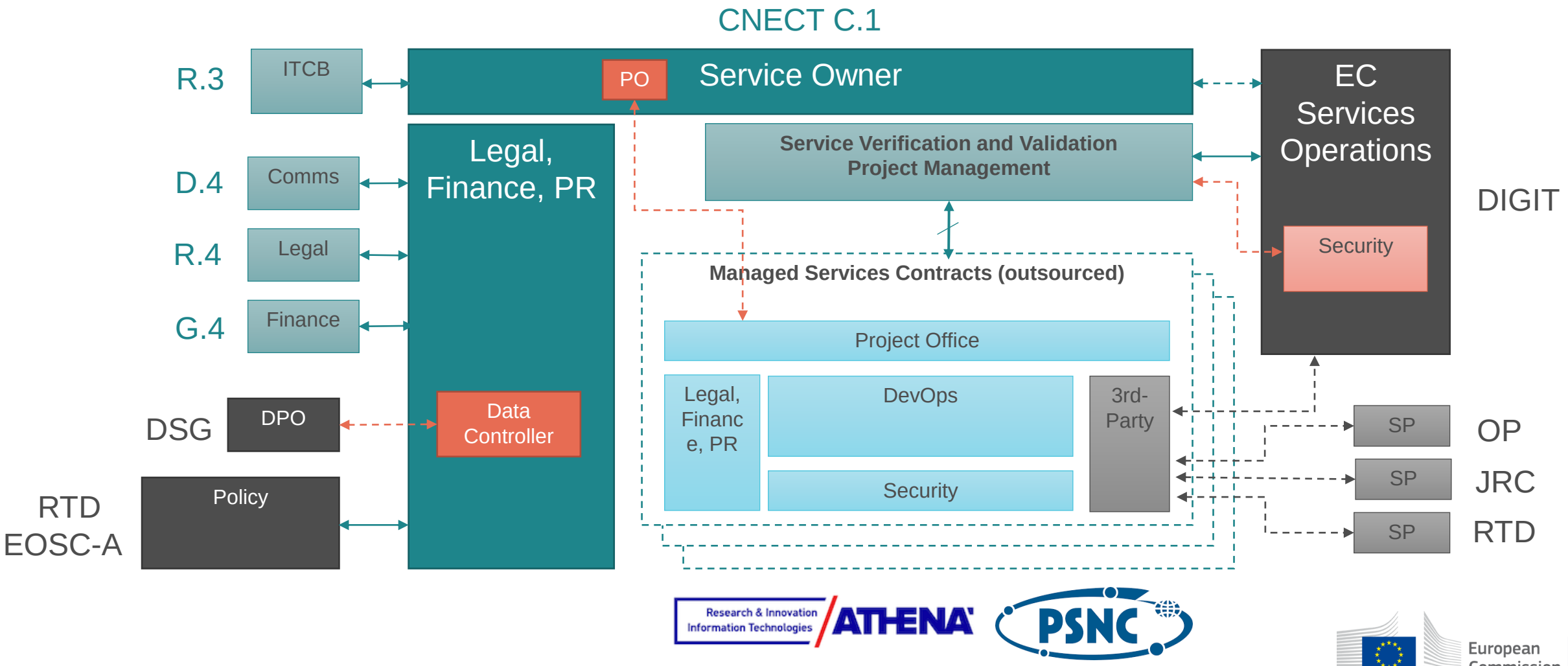
netcompany
intrasoft

icm

aws

European Commission

Operational Structure



Research & Innovation Information Technologies ATHENA PSNC

European Commission

Planned Timeline and Milestones

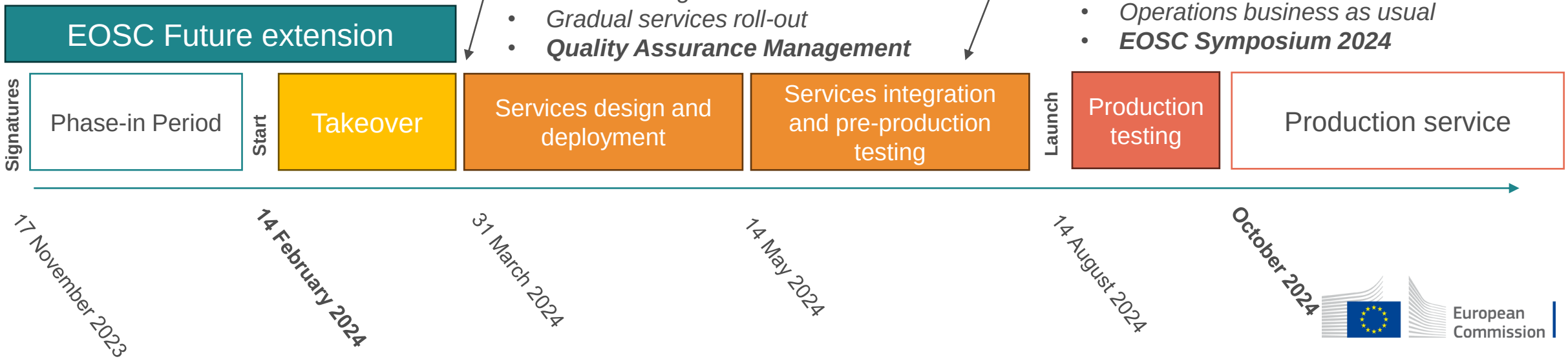
- *Planning and architecting of service components*
- ***Compliances and approvals (EC IT Gov)***
- *Definition of policies (RoP, AUP, AP)*
- *Production launch of the EOSC EU Node web front-office*

- *Full integration testing of core and exchange services*
- *Integration of EC services (ORE, ODP, CORDIS/Cellar)*
- ***Onboarding of flagships to EOSC EU Node***

- *Handover of relevant EOSC Portal functionalities (specs, software, IPO, etc.)*
- *Managing expectations of all stakeholders (providers, users, etc.)*

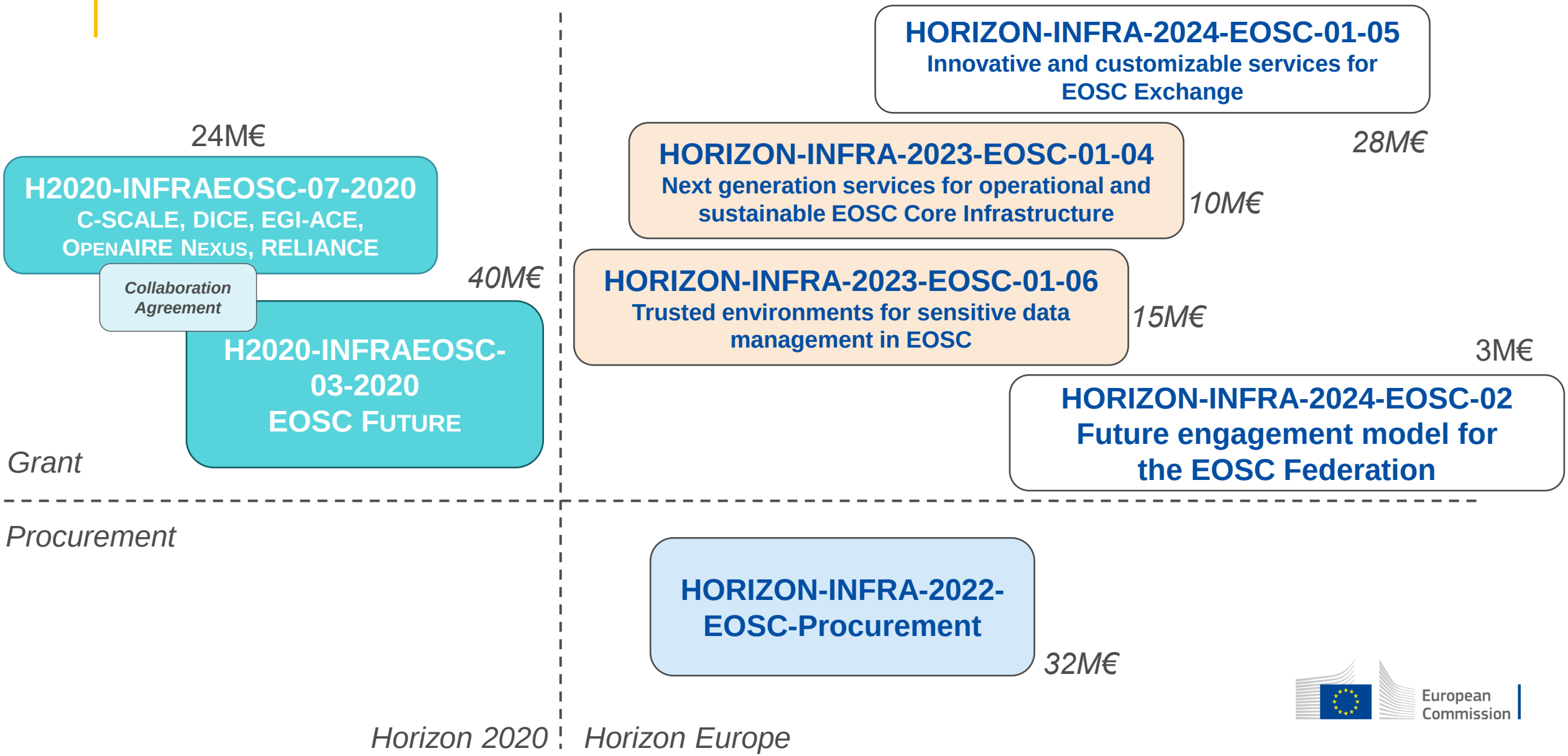
- *Deployments and staging of core/exchange functionalities*
- *Gradual services roll-out*
- ***Quality Assurance Management***

- *Pre-production and production testing*
- *First user experiences*
- *Operations business as usual*
- ***EOSC Symposium 2024***



European Commission

EU investments in direct support to the **EOSC Infrastructure**



European Commission

Other EU investments through the INFRAEOSC Destination

Enabling Open Science

Supporting an EOSC-ready digitally skilled workforce

HORIZON-INFRA-2021-EOSC-01-01
SKILLS4EOSC (7Mio€)

Services that underpin a research assessment system that incentivises Open Science

HORIZON-INFRA-2022-EOSC-01-01
GRASPOS (8 Mio€)

Supporting institutional open access publishing across Europe

HORIZON-INFRA-2022-EOSC-01-02
CRAFTOA (5 Mio€)

FAIR implementation

Deploying EOSC-Core components for FAIR

HORIZON-INFRA-2021-EOSC-01-03
FAIRCORE4EOSC (10 Mio€)

Enabling discovery and interoperability of research objects across communities

HORIZON-INFRA-2021-EOSC-01-05
FAIR-IMPACT (10 Mio€)

Support to international standards and specififcations for open sharing of FAIR research digital objects

HORIZON-INFRA-2022-EOSC-01-04
RDA TIGER (3 Mio€)

Planning, tracking and assessing scientific knowledge production

HORIZON-INFRA-2023-EOSC-01-03 (8 Mio€)

Improving the quality of scientific software and codes

HORIZON-INFRA-2023-EOSC-01-02 (8 Mio€)

Long term access and preservation infrastructures and data quality
HORIZON-INFRA-2024-EOSC-01-04 (8 Mio€)

Enabling a network of EOSC federated and trustworthy repositories
HORIZON-INFRA-2024-EOSC-01-03 (5 Mio€)

Uptake – Use cases

FAIR and open data sharing in support of the Cancer Mission

HORIZON-INFRA-2021-EOSC-01-06
EOSC4CANCER (8Mio€)

FAIR and open data sharing in support of the Mission on oceans & waters

HORIZON-INFRA-2022-EOSC-01-03
BLUE CLOUD 2026 AND AQUAINFRA (16 Mio€)

Build on the science cluster approach to ensure EOSC uptake

HORIZON-INFRA-2023-EOSC-01-01
(25 Mio€)

FAIR and open data sharing in support of the Mission climate adaptation

HORIZON-INFRA-2024-EOSC-01-01
(16 Mio€)

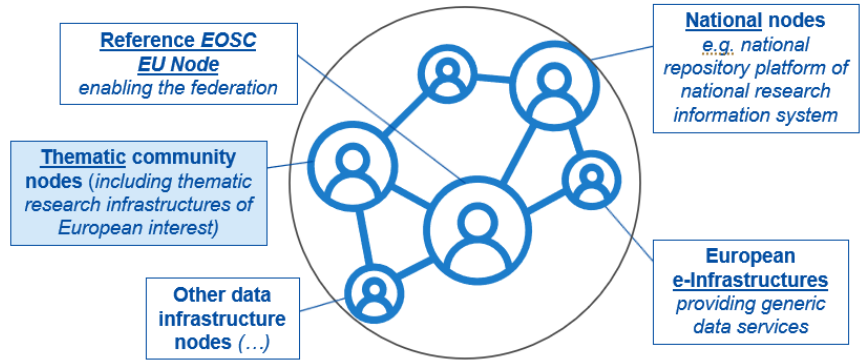
EOSC partnership

Supporting activities of the European EOSC Partnership

HORIZON-INFRA-2021-EOSC-01-02
EOSC-FOCUS
(4Mio€)

Supporting activities of the European EOSC Partnership
HORIZON-INFRA-2024-EOSC-01-02
(4 Mio€)

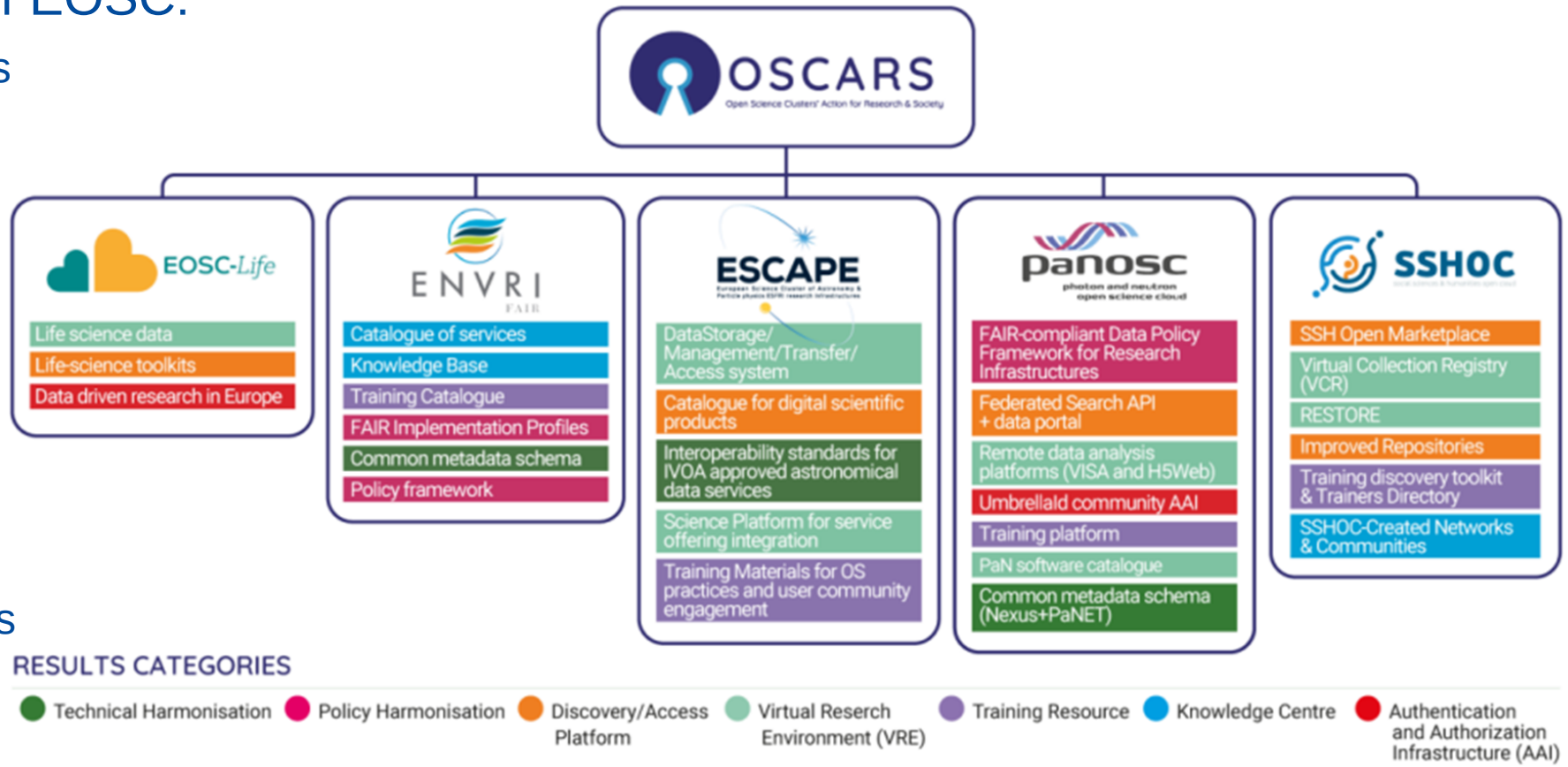
Thematic community Nodes: Some candidates



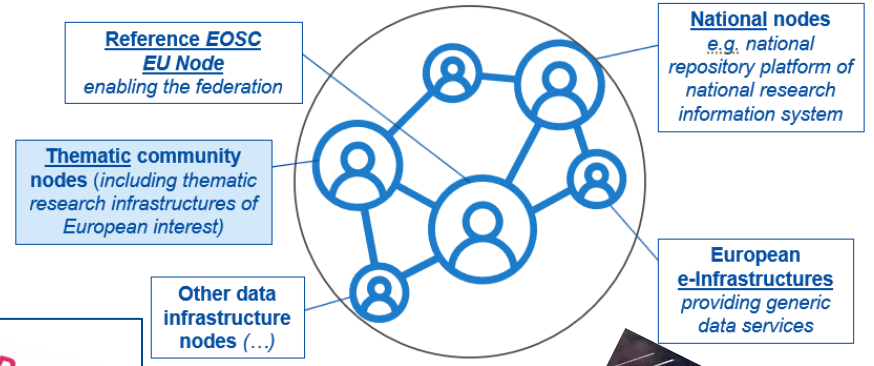
The Science Clusters approach:

Bottom-up implementation of the cross-border, cross-disciplinary model of EOSC:

- In H2020: from individual RIs to clustered RIs within 5 scientific domains (with EOSC onboarding)
- In HE: from a domain to a cross-domain approach with connection to the EOSC Federation
- **More than 40 RIs involved in the 5 Science Clusters**
- Need to act at different levels to address both specialization and generalization



Thematic community Nodes: Other candidates



The Blue-Cloud infrastructure:

EOSC blueprint for oceano-graphic research. More than 10 million data sets; about 1500 users per month, between 1000 and 3000 working sessions by individual users per month.

- FAIR data lake with central catalogue and common discovery and access service;
- Virtual Research Environment with storage and analysis capacity;
- 6 Virtual Labs to address scientific questions.

The European COVID-19 Data Portal:

Launched as an EOSC pilot in April 2020.

Over 25 million COVID-related, FAIR data records accessed by over 300.000 users in 187 countries.

The Pathogens Portal

Extends (since July 2023) to more than 200.000 pathogen species.

New research data commons under development

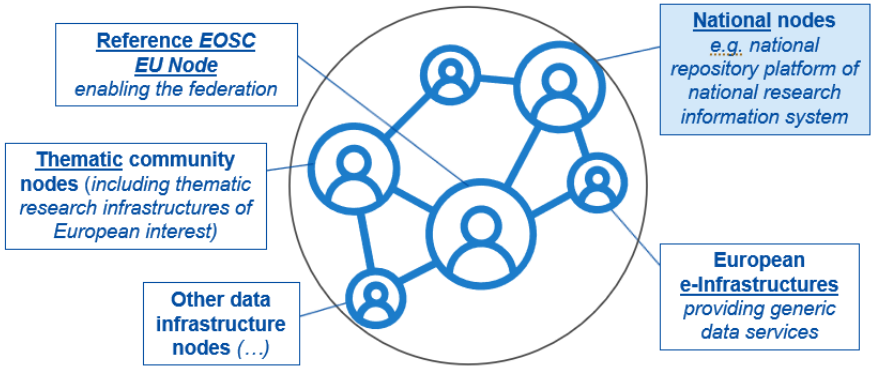
e.g. The European Collaborative Cultural Heritage Cloud or the “Materials Commons”.

Data Terra: a French infrastructure with international outreach

The Data Terra research infrastructure offers services relating to Earth system data that are interoperable and inter-disciplinary at all levels.

- Data discovery and access
- Production and data exploitation
- On-demand analytics and processing.

Candidate National Nodes



EOSC European co-programmed partnership:
In-kind contributions by non-EU partners exceeding **80 Mio€ per year** to upgrade existing research infrastructures and e-infrastructures so that they may be **federated through EOSC**.

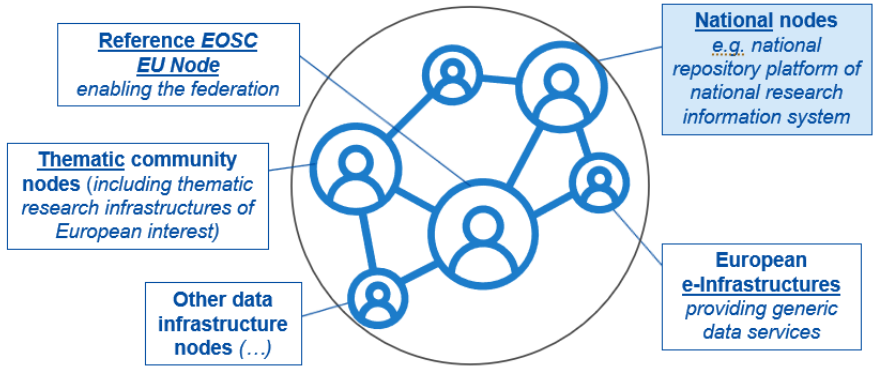
Open Science infrastructures in France

- Open access: **HAL** is a platform to promote Open Access to publications. Publications are easy to find, well referenced by search engines and interconnected with other services (ORCID, preprint servers).
- Open source: **Software Heritage** collects, preserves, and shares software that is publicly available in source code form.
- Open data: **Recherche Data Gouv** provides a repository (with Core Trust Seal certification) to deposit and disseminate data and a registry to search for data published in the repository itself or other external repositories. It aims to become an EOSC service.

Open Science infrastructures in Croatia

- The **Portal of Croatian scientific and Professional Journals** (HRČAK) includes 530 OA journals and provides access to 270,000 OA papers.
- The **Digital Academic Archives and Repositories (DABAR)** currently hosts 159 repositories and 212.000+ digital objects.
- The **Isabella computer cluster** hosted by SRCE provides significant computer resources (**EOSC onboarded**).

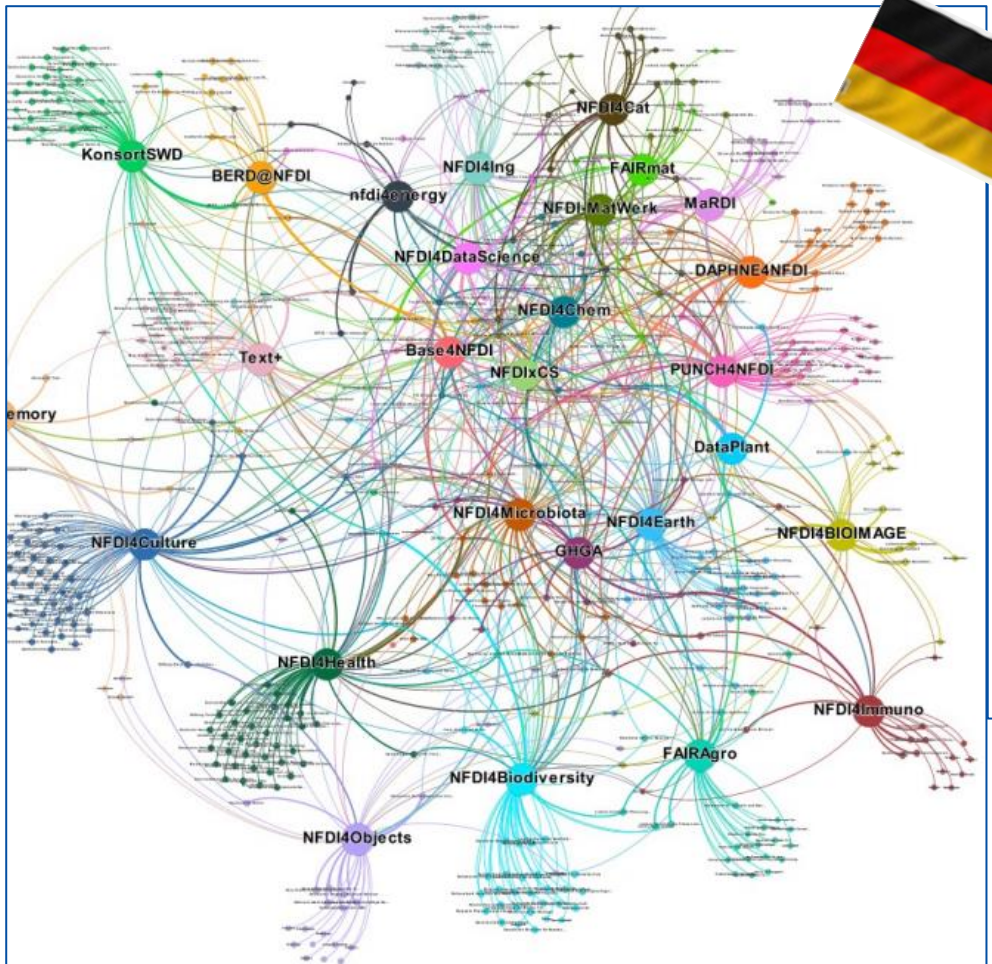
Candidate National Nodes



NDI: The National Data Infrastructure in Czech Republic

National Czech contribution to EOSC. It includes:

- The National Metadata Directory (NMA)
- The National Repository Platform (NRP)
- Thematic and possibly other (physical) repositories
- Policies, conditions of access, participation and use
- Training and educational activities (coordinated by the EOSC-CZ Training Centre)



NFDI: The German National Research Data Infrastructure:

27 NFDI consortia involving 261 association members are defining the basic infrastructure required at national level in Germany.

European Commission

Thank you

European Commission