



The European Open Science Cloud (and its role in the EU strategy for data)

Javier Lopez Albacete,

Open Science and Research Infrastructures

DG research & Innovation, EC.

13 March 2024

CERN

The European Open Science Cloud – EOSC

EOSC at the intersection of two EU priorities:

EU policy for Open Science

The **Open Science (OS) paradigm** affects the whole research cycle and all its stakeholders

It implies sharing knowledge and tools:

- “as early as possible” in the research process;
- “as openly as possible”;
- “as FAIR as possible”;

not only within a discipline but also between disciplines and society at large.

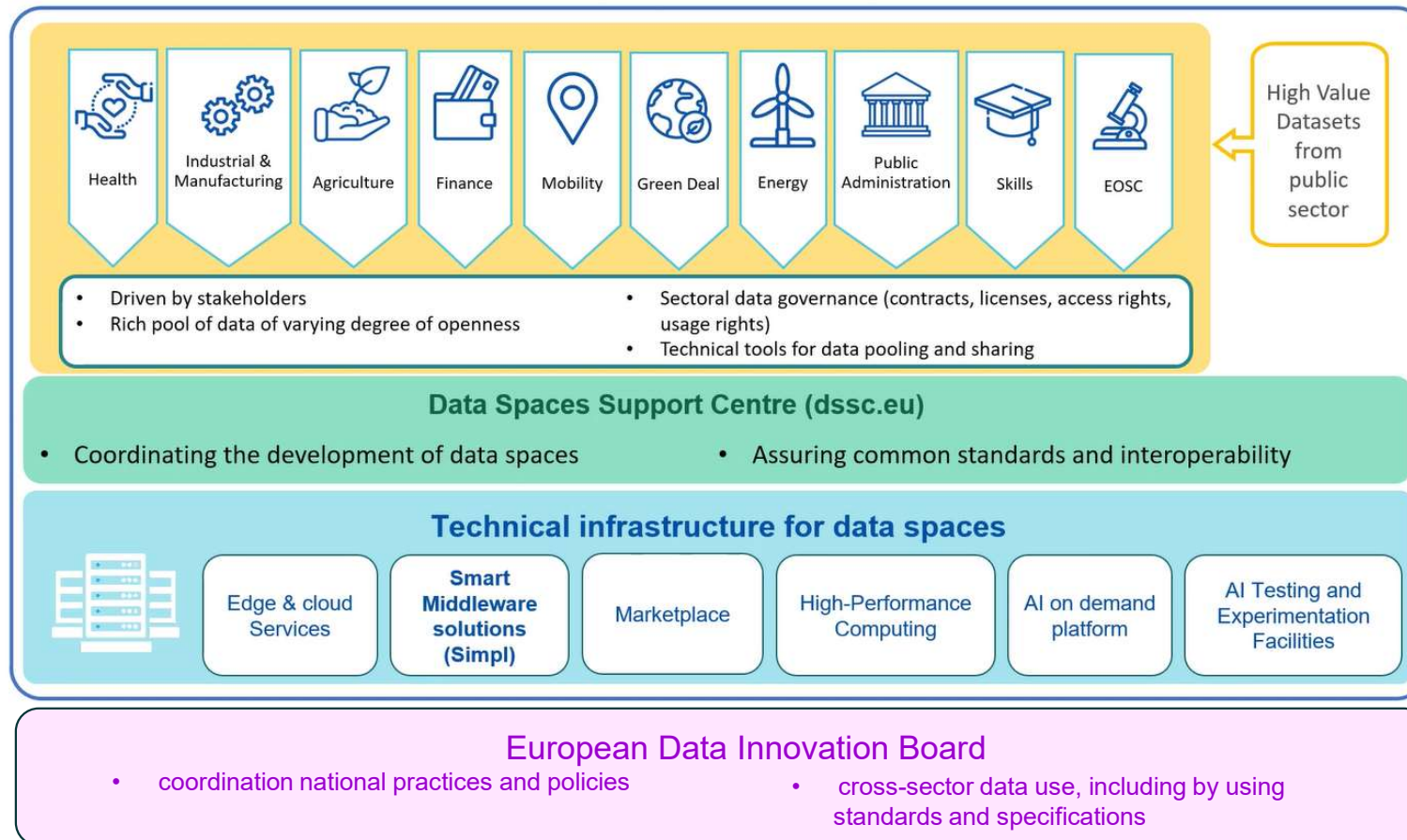
EU Strategy for data

Aims at creating a **single market for data** that will ensure Europe’s global competitiveness and data sovereignty

- Overcoming barriers to sharing, through technical **infrastructure**, **legal** rules and **ethical** guidelines.
- Fostering the development of collections of sector-specific data.
- Ensuring EU autonomy in supplying European cloud services.

Common European data spaces

Source: DG CNECT



Additional data spaces have been created for:

- Cultural heritage
- Language
- Media
- Tourism

And more could follow

Simpl system

- **Middleware** enabling **cloud-to-edge federations** and supporting common **European data spaces**
- “**Software glue**” ensuring **interoperability** within and between **data spaces**
- **Aim** to provide a **generic system** to enhance **interoperability** and **cost-efficiency**
- **Open-source** middleware allowing insights and simple deployment
- Funded via **Digital Europe Work Programme**
- **Open testing environment** planned to be launched in **early 2024**

Source: The European Commission's [website](#) on Simpl

Procurement lots

Simpl-Open

The core product of Simpl anchored in the preparatory study

Simpl-Labs

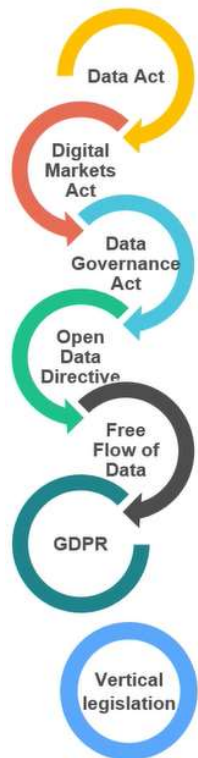
= playground and demonstration environment for Simpl-Open

Simpl-Live

= instances of Simpl-Open for sectoral data spaces, **including EOOSC**

EU data legislation

Source: DG CNECT



Aim	Data Covered	Regulated Actors
Ensure FAIRNESS in the allocation of data value among the actors of the data economy	Private sector data, personal and non-personal data, and co-generated (IoT) data	Businesses, public sector bodies, cloud and other data processing service providers
Tackle imbalances caused by the MARKET POWER of gatekeepers	Personal data and private sector data held by online platforms and originating from the users	Cloud and other data processing service providers, large data platforms
Ensure TRUST in data transactions	Public and private non-personal data, and personal data voluntarily made available by data holders	Data intermediation service providers, public sector bodies, (Recognised) Data Altruism Organisations
Promote use of OPEN DATA	Data in an open format that can be freely used, re-used and shared by anyone for any purpose	Public sector bodies, bodies governed by public law, public undertakings, universities
Ensure FREE FLOW OF DATA other than personal data within the Union	Non-personal data	Member States, competent authorities, professional users
Ensures a high-level of DATA PROTECTION and free flow of personal data in the Union	Personal data	Data controller, data processor, data subject, DPO, supervisory authorities, EDPB
Promote a competitive market according to SECTOR-SPECIFIC rules where necessary, e.g. automotive	Personal and non-personal data	Individuals and private and public sector bodies

Exceptions for re-use with research purposes under exploration through ERA Action 2

The European Open Science Cloud

Europe's contribution to the Web of FAIR data

A process

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

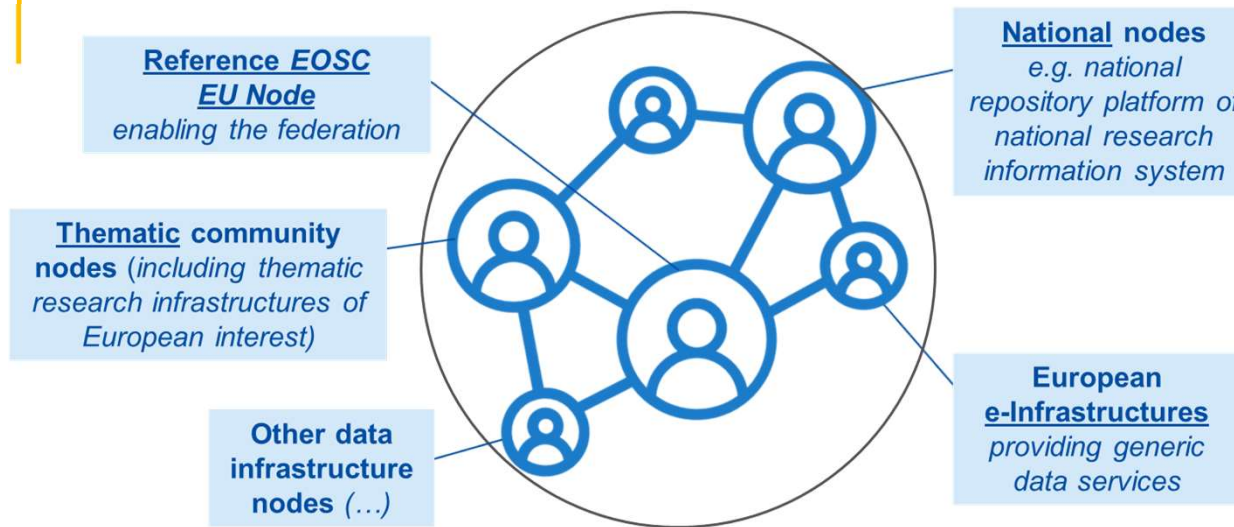
An evolving ecosystem

- 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.

An open, trusted, federated infrastructure

- To access existing Research Infrastructures in Europe;
- To enable over 2 million European researchers to store, share, process, analyse, and reuse research digital objects (e.g., data, publications and software)

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)

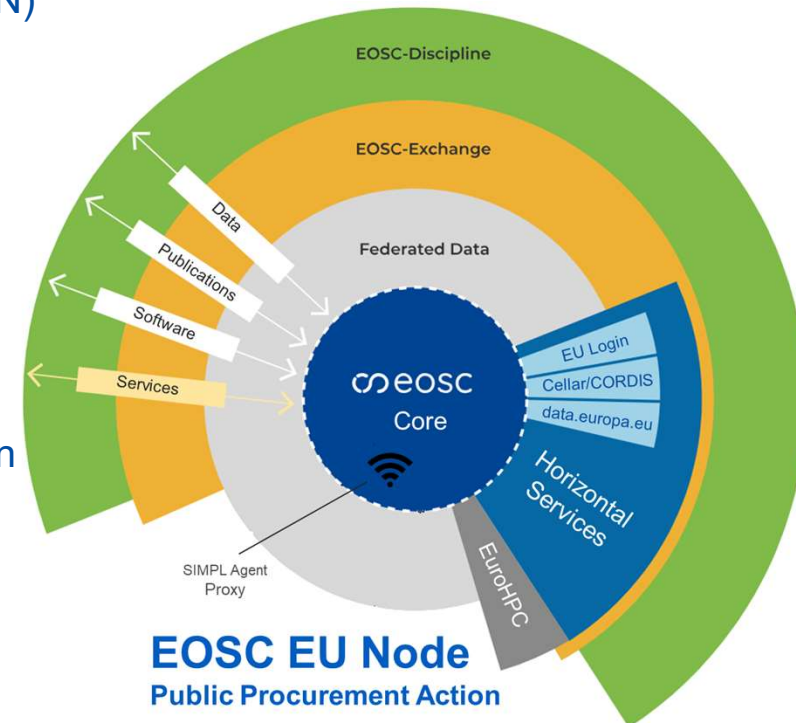
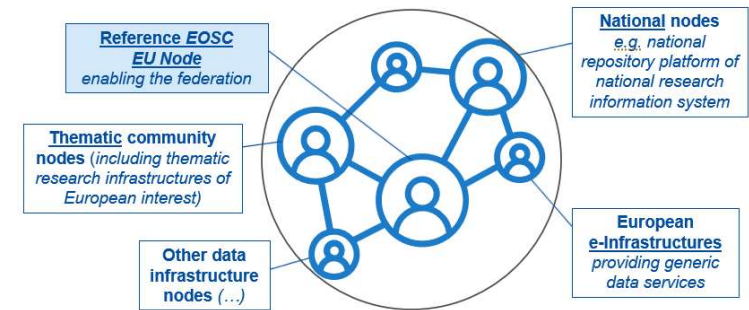
EOSC Node*: Data infrastructure system of variable nature (national, regional, institutional or thematic) with consensus-based policies, transparent ownership and clear responsibility, connected to the EOSC Federation to share information and resources within the EOSC community and to leverage common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

The EOSC EU Node

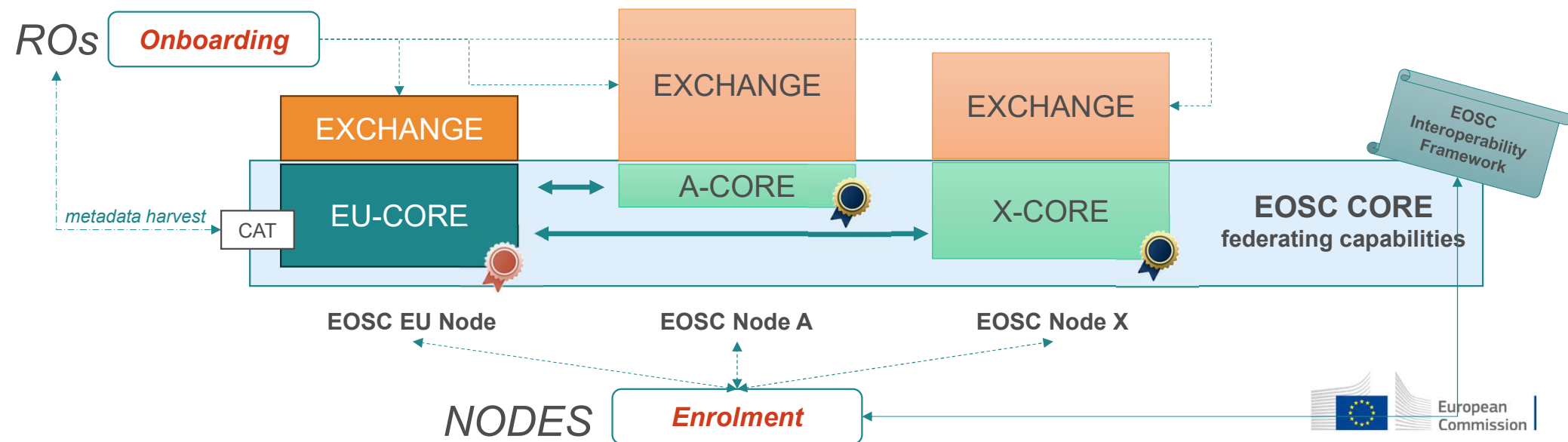
- **Puts** a “seed in the ground” by operating 24/7 the first recognised EOSC Node at the European level for an initial period of 3 years.
 - European level **multi-disciplinary** and **multi-national** scientific data/service portfolio for all research users (eduGAIN) and citizen scientists (EU Login/eIDAS)
 - **Offers “federating capabilities”**:
 - Identity Management
 - Resource Catalogues and Registry services
 - Application Workflow Management
 - Service Monitoring and Accounting
 - Management System (incl. Helpdesk)
- and common “horizontal services” for end-users to benefit from (compute, containers, data transfer, notebooks, file sharing, open research data)
- **SIMPL Agent proxy** to connect to other Data Spaces
 - **32 Mio €** from RI WP 2022 to cover three years of operation



EOSC Nodes in the federation

- What it takes to be an EOSC Node?

- **Technical capabilities** – Participate and contribute to the EOSC CORE federating capabilities (deploy, operate and manage at least one EOSC CORE capability, apply the EOSC Interoperability Framework)
- **Legal responsibilities** – Established entity under a recognized legal structure (e.g., JU, ERIC/EDIC, Governmental Org, Inter-governmental Org, International Org., etc.)
- **Administrative ownership** – Take ownership and participate in the EOSC Governance (e.g., EOSC Association)



The connections in the Federation

To address the EOSC puzzle

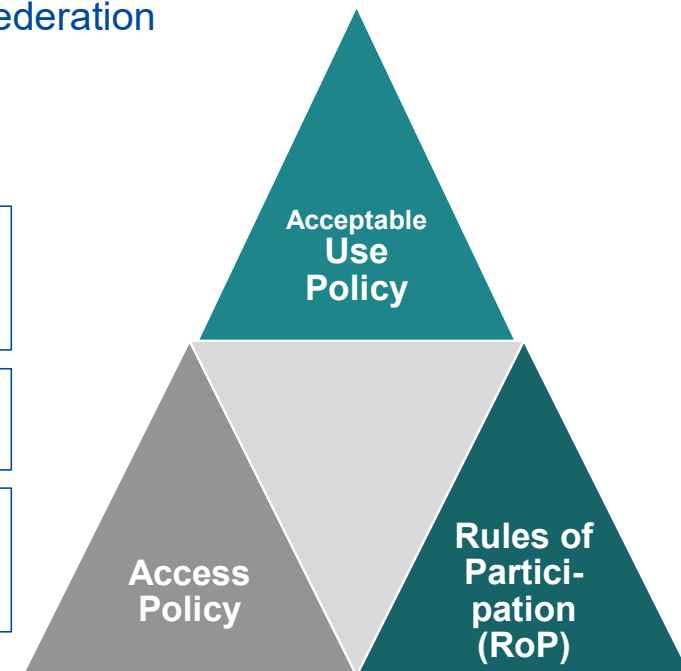
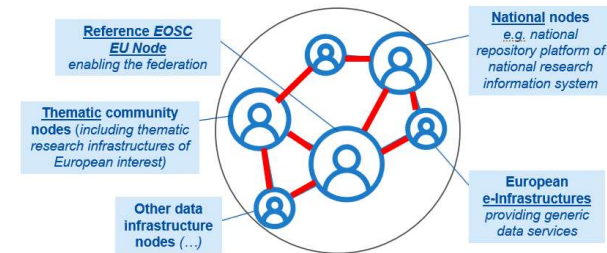
And cope with a vast quantity of infrastructure components of various scales and scopes, centralised or distributed, being generic, domain-specific or cross-disciplinary at institutional, national or European level.

- Minimal standards to be applied across the Federation (EOSC Interoperability Framework).
- Common policies and Rules of Participation applicable across the Federation
- A model to govern, enforce and evolve those policies.
- Trade-off between inclusiveness and quality of EOSC resources ?

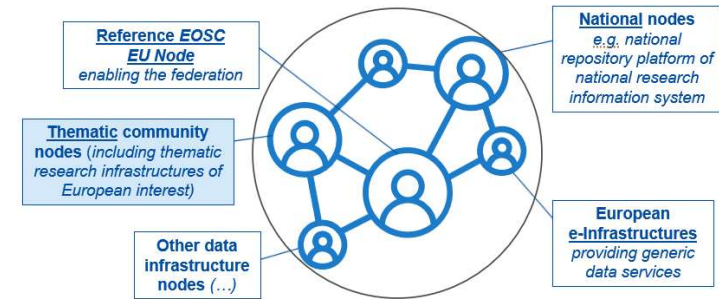
The Use Policy affect **all users** (registered or anonymous) and defines what is acceptable use of EOSC. For example: lawful, respect of intellectual property, respect of security rules...

The Access Policy primarily regulates the use of resources (computing, data, network) and security rules **for registered users**.

The Rules of Participation affect all **resource providers / Node operators**. Define the **minimal standards** for any 3rd party service (ie: assets, support) provided within the Federation. Deeply related to onboarding.



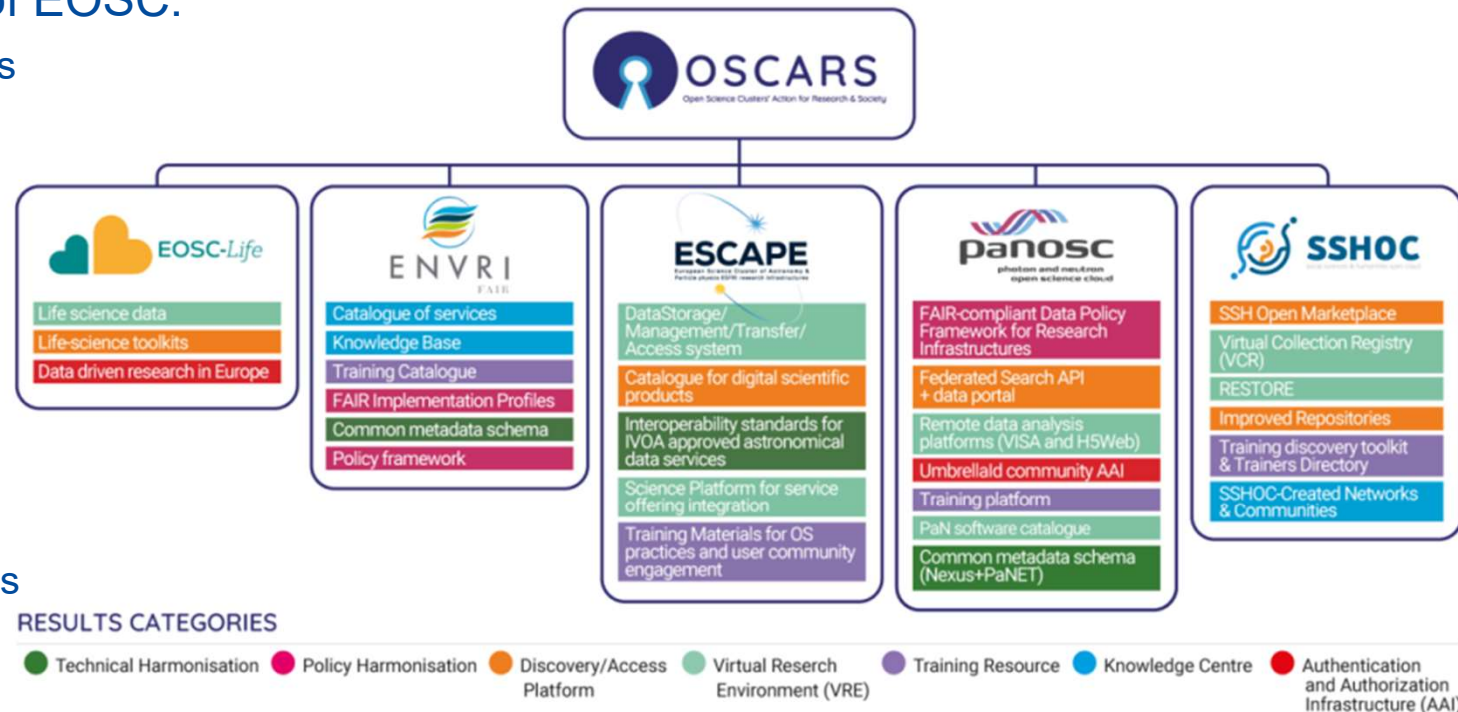
Thematic community Nodes: Some candidates



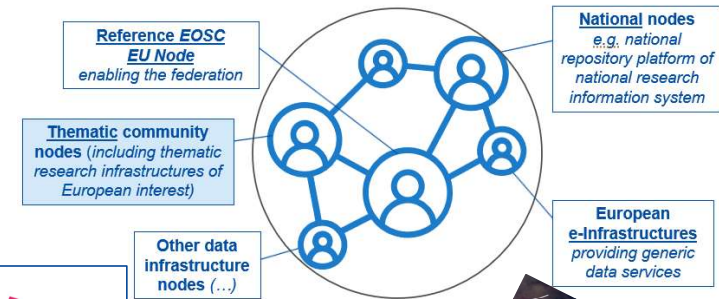
The Science Clusters approach:

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 oceanographic 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.

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.

EOSC in the countries

Increasing number of MS/AC have national policies/legislation in support of OS, as well as implementation strategies and funding mechanisms, thereby strongly increasing the EOSC readiness at the national level.

eOSC OBSERVATORY

23 out of **40**

countries have a national policy on open access publications

10 out of **40**

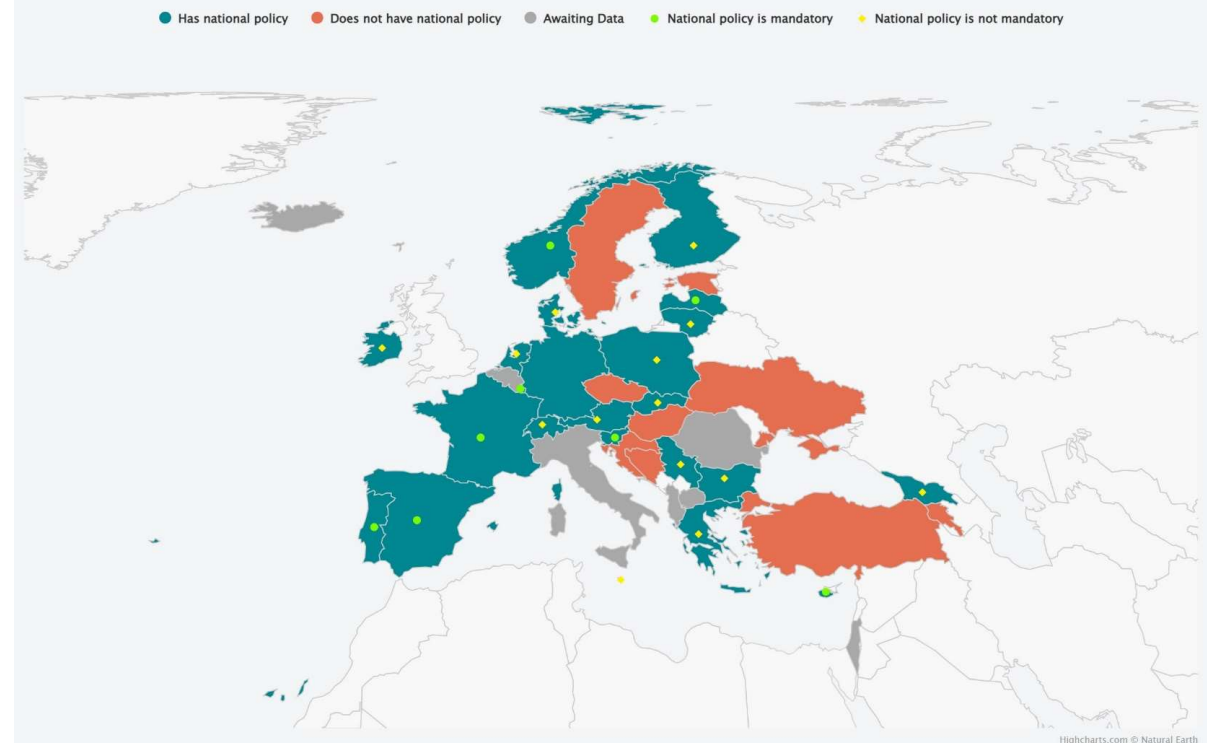
countries have a national policy on data

6 out of **40**

countries have a national policy on open source software

6 out of **40**

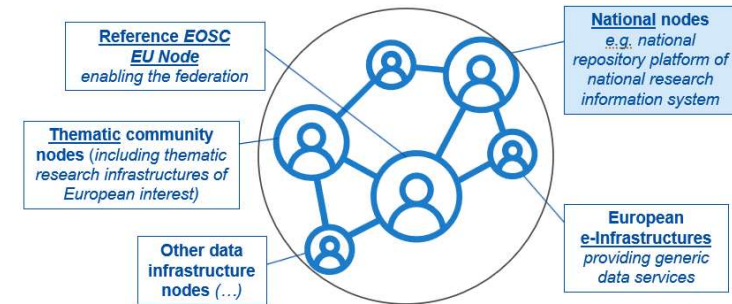
countries have a national policy on offering services through EOSC



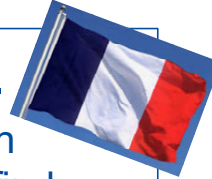
(Possible) National Nodes

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



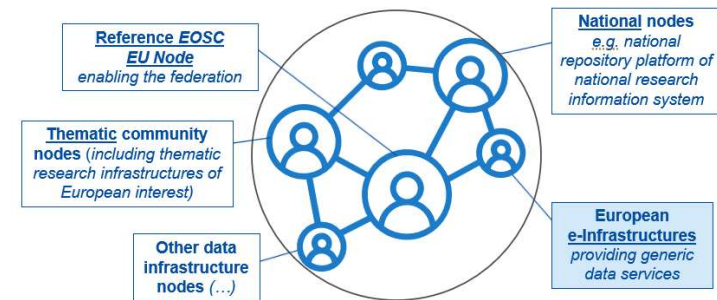
- Open access: **HAL** is a platform to promote Open Access to publications. Publications are easy to find, well referenced by search engines and interconnected with other services (ORCID, preprint servers).
- Open source: **Software Heritage** collects, preserves, and shares software that is publicly 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**).

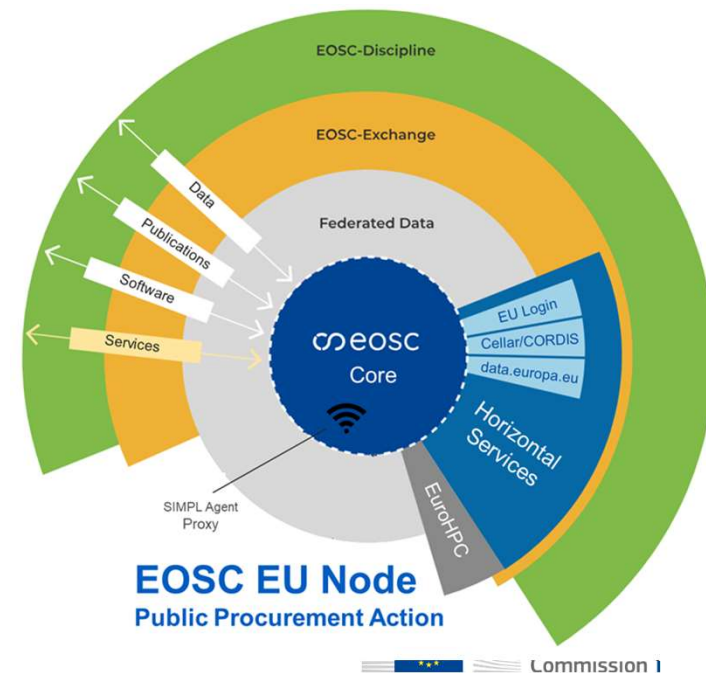
European e-Infrastructure Nodes



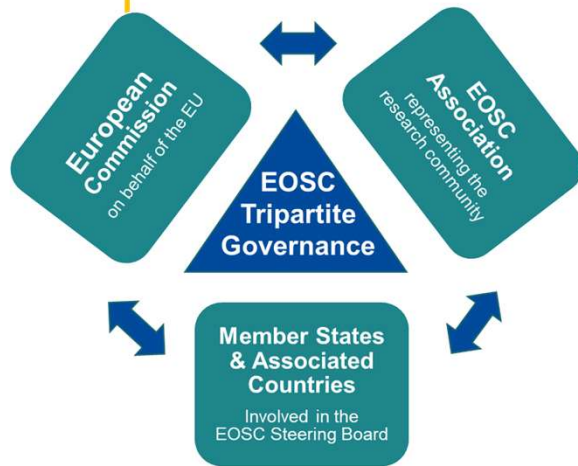
Multiple paths to integrate European e-service providers (pan-EU, national or private) in the EOSC Federation:

- As part of a consortium supporting directly the EOSC Core;
- As providers of some central services provided by the EU-Node to end-users of the Federation;
- As providers of services exposed on EOSC Exchange;
- As providers of e-Science services to specific ESFRI RIs, ERICs or other thematic Nodes connected to the EOSC Federation.

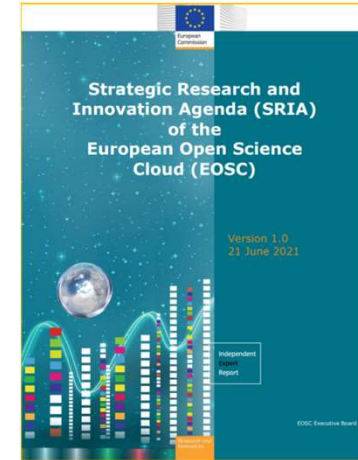
EOSC aim is to progressively to expand its user base to include the wider public sector and the **private sector**



EOSC Governance 2021-2027



- **EOSC tripartite governance** to ensure dialogue and strategic coordination between the European Commission, the Member States and Associated Countries and the EOSC Association
- **EOSC European co-programmed Partnership** to pool commitments and resources along priorities set in the **EOSC Strategic Research and Innovation Agenda**.
1 Billion € euro joint investment for 2021-2030.



Gap Analysis and some questions for post 2027:

- What will be the form of the future EOSC governance?
- What level of integration (policies, investments, services) are aimed at?
- How to ensure sustainability and evolve and improve the EOSC value proposition?
- How to ensure effective scalability of the Federation and to enable and adapt dynamically a series of business models?
- How are the different tasks entailed by EOSC to be shared among stakeholders?
- Who should operate the EOSC EU node?
- How to improve FAIR data productivity across for data intensive research?

The INFRAEOSC Destination in Horizon Europe: 2021 - 24

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€)

Long term access and preservation infrastructures and data quality

HORIZON-INFRA-2024-EOSC-01-04 (8 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 specifications 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€)

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€)

Thank you



© European Union 2023

Unless otherwise noted the reuse of this presentation is authorised under the [CC BY 4.0](https://creativecommons.org/licenses/by/4.0/) license. For any use or reproduction of elements that are not owned by the EU, permission may need to be sought directly from the respective right holders.

