



## EOSC Multi-Annual Roadmap 2023-24

**This document provides the EOSC Multi-Annual Roadmap for 2023-24. It has been developed by the EOSC Association Board of Directors and EOSC Association Task Forces. It is being released for consultation with the wider community in March 2022. Post consultation, this text will be incorporated into the Strategic Research and Innovation Agenda (SRIA) for EOSC as section 8 and a new version will be issued.**

### 1 Roadmap

This Strategic Research and Innovation Agenda (SRIA) sets clear goals to develop EOSC and build a research environment that promotes Open Science and increases trust in and reproducibility of research outcomes. The overall impact is a pan-European research landscape that offers significantly improved discovery, access, interoperability and exploitation of research outputs for researchers and for research and innovation stakeholders.

The Horizon Europe (HE) funding periods are mirrored in the SRIA implementation stages. Each stage defines a high-level objective for the period, aligned with the vision of enabling a trusted, virtual, federated environment in Europe to store, share and reuse research outputs across borders and scientific disciplines. For each stage, priority areas are defined, their prioritisation and scheduling reflecting community feedback from the open consultation process.

#### **Stage 1 (2021–2022): Development towards added value from a functional federation of infrastructures**

Enabling the **European Open Science Cloud operations (the EOSC-Core)** to provide necessary core functions of the Minimum Viable EOSC (MVE) that allow federation of existing and future infrastructures, according to associated Rules of Participation and governance that provision growth and expansion in the following stages.

#### **Stage 2 (2023–2024): Expansion to production that generates added value**

Expanding and building the core data infrastructure to support the full cycle of workflows for scientific research in key thematic areas. During this period, activities will build on pilots/demonstrators and work towards linking EOSC beyond the research communities to the wider public sector and the private sector will begin.

#### **Stage 3 (2025–2027 and beyond): Expansion to develop impact from Open Science**

Deployment of federated research infrastructures for European researchers with functionality that provisions actors from multiple communities to deliver impactful Open Science. In addition to European infrastructures, the national research infrastructures delivered from the Member States and Associated Countries in particular will help in this expansion phase.



This section proposes a work plan for the 2023-24 period, taking into account the activities delivered in the first stage and those forthcoming under the Horizon Europe INFRAEOSC Work Programme. It has been drafted by the EOSC Association Board of Directors in collaboration with the EOSC Association's Task Forces. The content has also been shaped by a community consultation held in March 2022, which provided inputs on priorities. The roadmap provides a framework of priorities with derived activities and indicators for the second stage of activity. These focus on embedding EOSC by increasing the number of research infrastructures that are federated and growing the range of research outputs and services that are available. Activities also focus on developing skill sets, offering support networks and increasing the adoption of standards so Open Science and FAIR practises are more widely adopted. The EOSC user base should also be widened through collaboration with other Partnerships, data spaces and public sector groups, and key priorities such as developing infrastructure for sensitive data sharing are intended to support the inclusion of data from public authorities.

### *1.1. Results of the MAR open consultation*

*To come post consultation in March*

### *1.2. Minimum Viable EOSC (MVE)*

A functioning EOSC can only emerge from a coordinated effort of H2020, Horizon Europe, national and institutional activities. The overall vision of the SRIA is the summation of activities at all three levels: European, national and institutional. Coordination of these activities is essential and must be well defined. The MVE shall deliver on the core SRIA objectives and provide foundational support for Open Science in Europe.

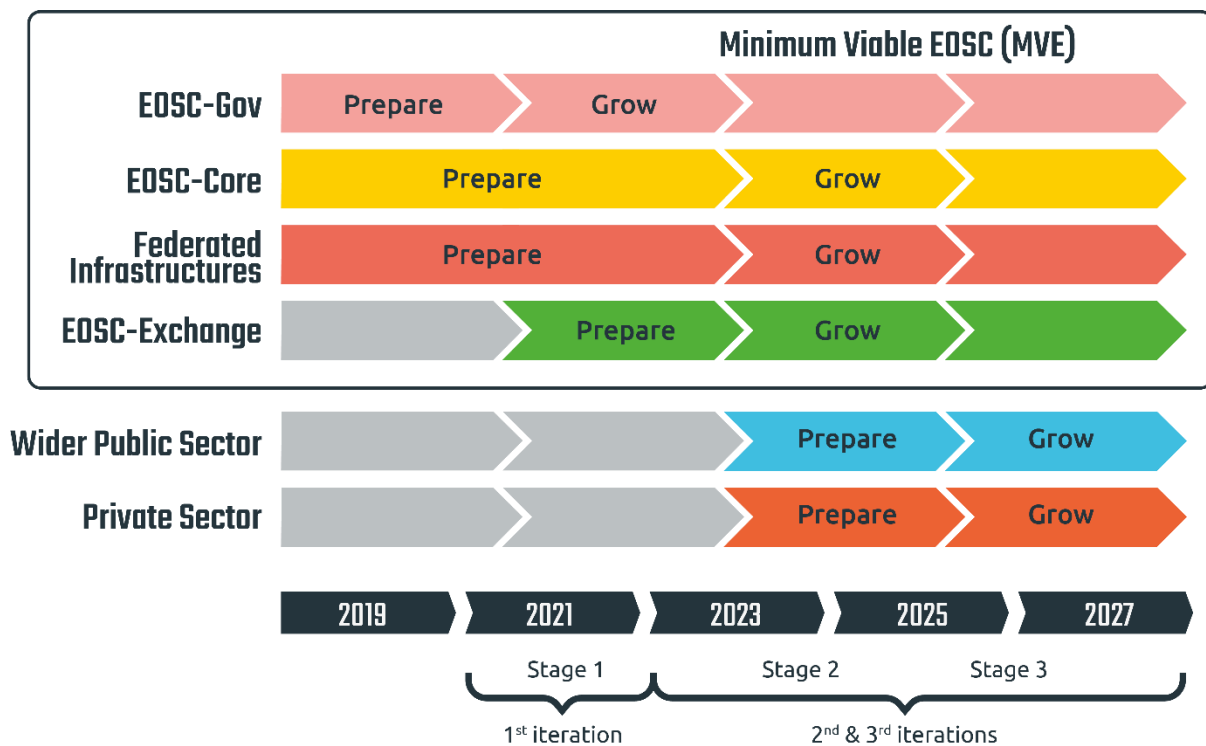
The number, complexity and interdependency of existing projects illustrates the ongoing challenge of building EOSC and raising the importance of Open Science within the research community to a point where Open Science is the new normal. Horizon Europe will build on the results of existing projects both for core functionality and for thematic activities centred around federated data infrastructures.

As the voice of the community, the EOSC Association plays a pivotal role here. It oversees the implementation of the strategic objectives defined in this SRIA and will take over results from the EOSC projects, where assessments performed by the community and the EOSC-A Board recommend they are fit for purpose. The EOSC Association will stimulate the coordination of activities to progress the MVE towards a functional and performant federated data infrastructure. For this to be effective, the EOSC Future project, which is the flagship infrastructure investment under Horizon 2020, and coming Horizon Europe projects will have to work on convergence together with the EOSC Association. For the coming projects this has to be written into their grant agreements, and the help of the European Commission will be key to achieve the right structure.

The EOSC Future project is tasked with developing an operational EOSC platform, adopting a system of systems approach which federates and connects existing infrastructure rather than creating a new monolithic structure. The project brings together the e-Infrastructures with all five research clusters (ENVRI-FAIR, EOSC Life, PANOSC, SSHOC and ESCAPE) to ensure the integration of data and resources from their communities into the EOSC Platform.

### 1.2.1. Scope and timing of the MVE

The staged approach to the development of EOSC described in the SRIA is presented in Figure 1.1. The approach covers the key required areas and development periods to enable EOSC engagement over a broad community, extending beyond researchers to the private sector. The period of development described in the SRIA covers two periods of European Commission (EC) funding: Horizon 2020 and Horizon Europe. It is expected that EOSC will be delivered through the activities of a large number of engaged stakeholders. Stage 1 describes the activities to be initiated in the first iteration of the MVE in 2021–2022, Stage 2 describes the second iteration in 2023–2024, and Stage 3 the third iteration in 2025–2027.



**Figure 1.1: Scope and timing of the MVE**

The EOSC governance has been well developed during the first stage with the establishment of the EOSC Association and tripartite collaboration with the EC and Member States and Associated Countries through the EOSC Steering Board. The EOSC Association General Assembly now comprises 233 organisations, of which 160 are full Members and 73 are Observers. These members come from a wide range of countries and organisations types, including funders, universities, national facilities, commercial providers and not-for-profit bodies. Such representation enables it to act as the voice of the community. Moreover, a Secretariat of six members of staff has been established, including a Secretary General and Chief Technology Officer. The team will grow to 10 members of staff in 2022. As such, the Association is well placed to stimulate coordination across the funded EOSC implementation projects and represent the stakeholder community when steering developments. Its thirteen Task Forces will play an instrumental role in setting the agenda and coordinating

developments in specific areas of implementation.<sup>1</sup> As EOSC continues to evolve, playing an active role in making it operational will require a further increase in Association staffing.

The development of EOSC Core, EOSC Exchange and the federation of Research Infrastructures is also underway through the EOSC Future project. The project has advanced the EOSC Architecture description laid out by previous Executive Board Working Groups and has been developing the building blocks such as Interoperability Framework and associated governance process to allow these procedures to be implemented. Building on previous project outcomes, work has advanced to allow data catalogues to be harvested using the OpenAIRE-PROVIDE guidelines and research graph technology is being used to connect up the existing service and resource catalogues.

Activities in the roadmap for 2023-25 focus on further growth to embed these governance and core EOSC platform structures, as well as to prepare to extend work to the public and private sectors.

### 1.2.2. H2020 and Horizon Europe EOSC projects

The implementation of EOSC is driven by a series of European Commission funded projects related to EOSC. Projects (still) active from the 2021 period are represented in Figure 1.2. Several of these were supported under the Horizon 2020 work programme, with those starting since 2022 being supported under Horizon Europe. The wide range and diversity of projects illuminates the challenge facing the EOSC Association in terms of coordination. Attempting to build a coherent infrastructure via a series of projects is problematic, especially when their accountability and evaluation is based on individual work plans rather than the overarching EOSC strategic agenda. The proposed grant amendments under Horizon Europe to ensure collaboration and coordination with the EOSC Association should hopefully alleviate this position, but realising EOSC will also require a willingness on the part of projects and the EC to be flexible and allow changes to proposed plans to avoid overlapping activities or divergent approaches.

The INFRAEOSC calls are of particular importance since they support the major investment in building the EOSC Platform and fostering the development of FAIR and Open Science practises. Within this programme, a series of grouped projects have been supported to represent the science clusters, regional initiatives and free-at-the-point-of-access services. These groupings and associated projects are listed below for reference. Coordinating projects such as the EOSCsecretariat and EOSC Future have established collaboration agreements with these groups to foster alignment of joint activities:

#### **INFRAEOSC 04-2018** – Connecting ESFRI infrastructures through Cluster projects

- Photon and Neutron Open Science Cloud (PANOSC)
- ENVironmental Research Infrastructures building Fair services Accessible for society, Innovation and Research (ENVRI-FAIR)
- Social Sciences & Humanities Open Cloud (SSHOC)
- European Science Cluster of Astronomy & Particle physics ESFRI research infrastructures (ESCAPE)
- EOSC-Life

#### **INFRAEOSC 05b-2018** – Coordination of EOSC-relevant national initiatives across Europe

- Coordination and harmonisation of national initiatives, infrastructures and data services in Central and Western Europe (EOSC Pillar)
- EOSC Nordic

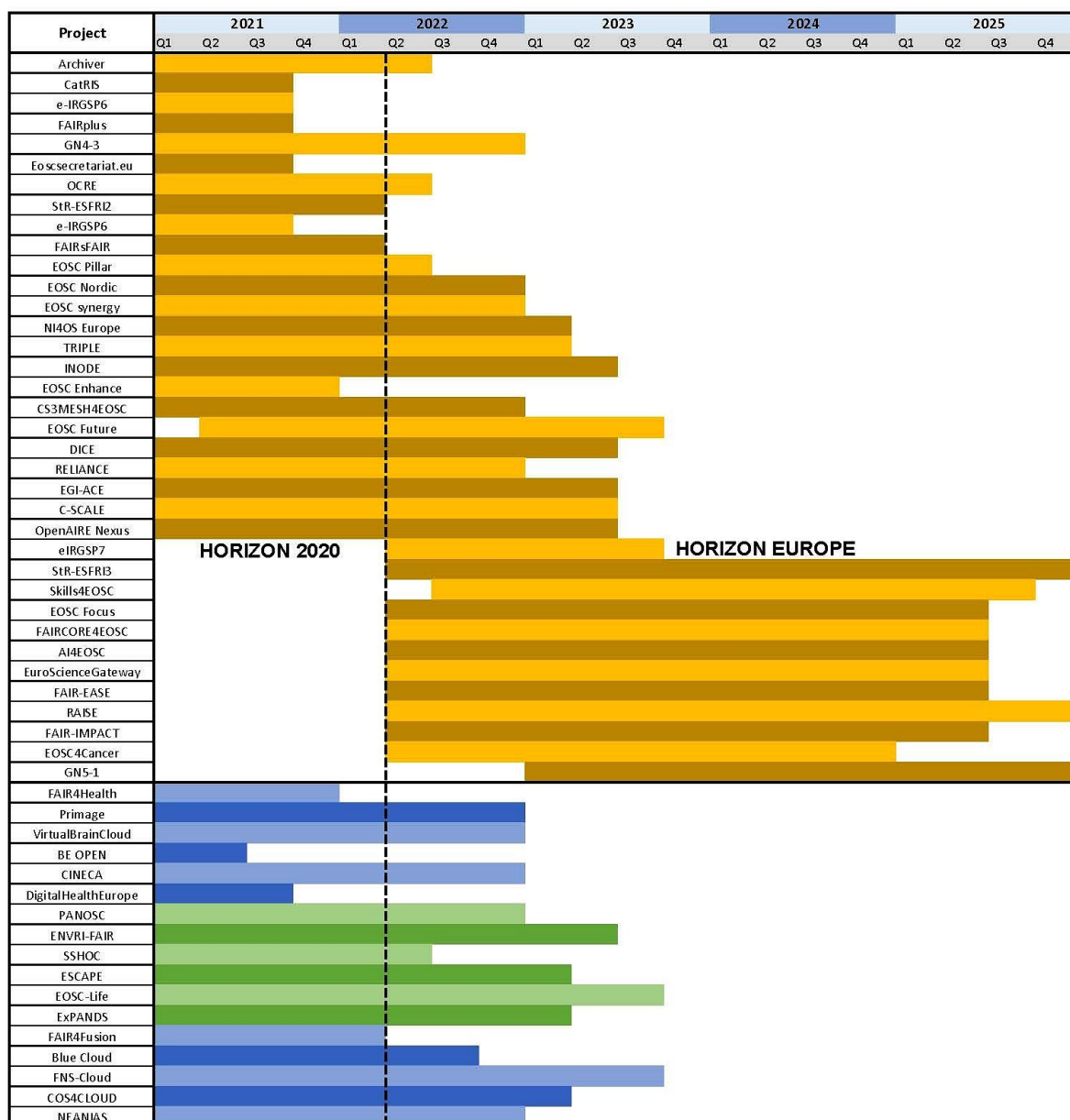
---

<sup>1</sup> EOSC Association Advisory Groups and Task Forces are listed at: <https://eosc.eu/advisory-groups>

- Expanding Capacities by building Capabilities (EOSC Synergy)
- National Initiatives for Open Science in Europe (NI4OS-EUROPE)
- EOSC Photon and Neutron Data Services (ExPaNDS)

**INFRAEOSC 07-2020** – Increasing the service offer of the EOSC Portal

- Data infrastructure capacity for the European Open Science Cloud (DICE)
- REsearch Lifecycle mAnagement for Earth Science Communities and CopErnicus users in EOSC (RELIANCE)
- EGI Advanced Computing for EOSC (EGI-ACE)
- Copernicus - eoSC AnaLytics Engine (C-SCALE)
- OpenAIRE Nexus Scholarly Communication Services for EOSC users (OpenAIRE-Nexus)



**Key:** Orange shades indicate coordination, best practice or technology projects. Blue shades indicate thematic science-based projects. Green shades indicate ESFRI cluster projects.

**Figure 1.2: EOSC projects active since 2021**

### 1.3. Objectives

The Co-programmed European Partnership on EOSC, as agreed between the partners, is to be implemented in an open, transparent, efficient and flexible way. The intended cooperative relationship aims to achieve jointly defined objectives based on a long-term common vision and a clear commitment from the partners throughout the duration of the Partnership. More on the EOSC Partnership can be found in section 2.9.1 of the SRIA.

Slight edits have been made to the Specific and Operational Objectives since version 1.0 of the SRIA. These include removing all dates, since only some objectives included timeframes and the delivery is tracked via the KPIs rather than objectives, and removing periphery text to assist in readability. The edits made did not change the meaning or emphasis of any objective. The only significant change made was to add Specific Objective (SO) 10. This relates to a new priority area of work, namely data quality, which was not reflected in the roadmap previously.

The **general objectives** (GOs) of the European Partnership, which are identical to the strategic objectives described in Section 3 of this SRIA, are defined as follows:

- GO1** Ensure that Open Science practises and skills are rewarded and taught, becoming the 'new normal';
- GO2** Enable the definition and adoption of standards, and the development of tools and services, to allow researchers to find, access, reuse and combine results;
- GO3** Establish a sustainable and federated infrastructure enabling open sharing of scientific results.

The **specific objectives** (SOs), which are reflected in the critical success factors identified in Section 7.4 of this SRIA, are the following:

- SO1** Increase in the number of relevant research outputs that are made available as open as possible by researchers performing publicly funded research;
- SO2** Professional data stewards are increasingly available in research-performing organisations in Europe to support Open Science;
- SO3** Development and adoption of incentives for researchers to perform Open Science;
- SO4** Increasing amounts of research outputs from publicly funded research in Europe are FAIR by design;
- SO5** The EOSC Interoperability Framework supports an increasing range and quantity of FAIR digital objects including data, software and other research artefacts;
- SO6** Provide an increased number of services and resources to ensure that European research is discovered and reused within and across disciplines to extract new knowledge;
- SO7** EOSC is operationalised and provides a stable and valuable infrastructure supporting researchers to address societal challenges;
- SO8** Essential additional functionalities for end users from the public and private sectors are implemented in EOSC (these developments are complementary to those of other European data spaces);
- SO9** EOSC increasingly establishes ties with related initiatives from regions around the world and becomes a partner in global cooperation frameworks for Open Science.
- SO10** Common data quality indicators are agreed and implemented to ensure that research outputs within EOSC are ready for FAIR usage.



The **operational objectives** (OOs), which are reflected in the action areas described in Sections 5 and 6 of this SRIA, are the following:

- OO1** Deliver and operate all the necessary components of the Minimum Viable EOSC to openly share research data, publications, software, tools and services while attracting increasing numbers and categories of users, including those from public and private sectors.
- OO2** Develop and evolve monitoring systems to gather data and evidence on best Open Science practises accessible through EOSC. This could utilise the EOSC Observatory dashboard to monitor the evolving landscape of policies, infrastructures and open resources;
- OO3** Increasingly mainstream Open Science skills in European research-performing organisations (RPOs), including through the uptake and alignment of curricula and training frameworks related to the full lifecycle of data management;
- OO4** Co-develop domain-specific standards and adopt Open Science practises through engagement with research communities;
- OO5** Provide the technical components of a FAIR ecosystem for uptake and customisation by the communities (including open specifications, standards, schemas, application programming interfaces (APIs), metadata frameworks supporting FAIR digital objects and their automated processing);
- OO6** Provide the metrics and tools to measure the adoption of the FAIR principles for research outputs and provide frameworks to help in certifying that repository services enable FAIR;
- OO7** Co-develop a first generation of a robust pan-European network of infrastructures for software source code (including incentives for the effective documentation and sharing of research software);
- OO8** Co-design and adopt a rewards and recognition framework for FAIR and open science practises in research;
- OO9** Implement and evolve the EOSC Rules of Participation and onboarding process for EOSC providers and increase the number of service providers and services offered progressively;
- OO10** Deploy and operate an authentication and authorisation infrastructure (AAI) framework to manage user identity and access;
- OO11** Implement the EOSC persistent identifier (PID) policy and architecture, including the development of a global PID resolver;
- OO12** Co-design a minimum metadata framework and provide a common search mechanism to EOSC resources;
- OO13** Continuously monitor and promote the increased uptake of core services and EOSC resources, access to EOSC Exchange tools and services and ensure a feedback loop with the users;
- OO14** Define models for availability and costing of services across borders.

#### *1.4. Levels of implementation*

The activities for implementing all stages of this SRIA are outlined below (Section 1.5). Those carrying out these activities can best be seen as operating at three levels of implementation:



- **European level (L1):** everything done at the European level either by or financed through the Horizon Europe programme or other sources, be it an effort by the EOSC Association or by a research infrastructure or service organisation operating at a European level. This could also be achieved through the joint effort of countries, etc. In other words, every contribution to EOSC being worked on at the European level (irrespective of who pays).
- **National level (L2):** the same as above but then at the country level, i.e. activities in one of the Member States or countries associated to the Horizon Europe programme (MS/AC), as long as the activities contribute positively to the development of the EOSC ecosystem as described in this SRIA.
- **Institutional level (L3):** again, the same as above but then the activities at the level of the participating institutions (e.g. a university or other research-performing organisation), and again as long as the activities are aligned with the country strategy which should in turn be aligned with the European strategy as described in the SRIA.

The activities described in Section 8.5 below are grouped by the three levels.

### *1.5. Priorities (2023-24)*

The priorities for 2023-24 are drawn from the evolution of the 2021-22 roadmap, the forthcoming European Commission work programme and the EOSC Association workplan. They focus on embedding the foundational governance structures and EOSC technical infrastructure achieved in 2021-22. Priorities are assigned to three different levels, namely:

1. European
2. National
3. Institutional developments

Priorities are structured according to the three main objectives listed in the SRIA. For each objective a series of priority activities and expected outcomes is specified.

## **Objective 1**

**Ensure that Open Science practises and skills are rewarded and taught, becoming the ‘new normal’**

### **Priority activities:**

#### Level 1: European

- Align EU policy and legislation with Open Science principles and practises.
- Liaise with Horizon Europe projects, EOSC Future and other key EOSC projects such as those funded under the INFRAEOSC 07-2020 and regional INFRAEOSC-5B calls to coordinate activities jointly leading to a Minimum Viable EOSC
- Manage and evolve the EOSC Observatory to record investments, policies, research outputs, open science skills and infrastructure capacities being planned and delivered via Member States, Associated Countries and EOSC Association members
- Review possible EOSC onboarding procedures ensuring the EOSC Rules of Participation are appropriately reflected, processes are automated where possible and barriers to uptake are low
- Setup a coordinated technical support programme for users from different target groups (research collaborations, long tail of science, citizen science, policy makers and private organisations) to federate their digital environments with a wide range of data, models,





workflows, scientific software and services from national providers, horizontal and thematic research infrastructures and e-Infrastructures.

- Widen and deepen the possible EOSC user base via collaboration with European Partnerships, the HE missions, Research Infrastructures, science clusters, sectoral data spaces and Citizen Science groups
- Operate an effective stakeholder forum which provides ample opportunities for the community to learn about, be consulted on and engage in EOSC developments
- Focus on engaging research communities to increase their participation in EOSC through for example Marie Curie Alumni Association, the science clusters, EURODOC, YAE, networks of research performing organisations and international open infrastructure communities such as ORCID.
- Continue to support the development of curricula for Open Science and FAIR skills and align these at a European level to ensure a base-level standard is reached across the European research community
- Support the development of networks for data stewards, research software engineers and Open Science communities at European level
- Formulate measures for researcher engagement in quality Open Science to reward researchers and incentivise adoption of these practices.
- Enable coordination within and across disciplinary communities to support the creation and dissemination of shared domain specific standards and evaluation frameworks
- Support research to systematically follow up and evaluate the impact of new metrics and procedures for research assessment aiming to incentivise open science practises

#### Level 2: National

- Harmonise national Open Science monitoring and KPIs to feed into the EOSC Observatory activities
- Focus on engaging research communities to increase their participation in EOSC through e.g. RDA / GoFAIR National Nodes.
- Coordinate awareness raising and uptake of EOSC in collaboration with the designated EOSC Association Mandated Organisations or other EOSC national structures, national funders and research infrastructures to interlink the national OS ecosystems to EOSC
- Develop national policies and funding streams to support the development of EOSC and foster collaboration between the national implementations of EOSC and the Research Infrastructures to incentivise and encourage national RIs to federate into EOSC
- Support regional corporation learning from the INFRAEOSC-5b projects to strengthen the national engagements
- Incentivise and encourage smaller community and institutional infrastructures in their growth towards ensuring FAIR and long-term access
- Establish the cost of data management, data stewardship and maintenance of research outputs (including software) as eligible within national funding schemes
- Develop curricula for Open Science and FAIR data skills in line with aligned European curricula and encourage the inclusion of these as a core element in research programmes
- Support training of data stewards and research software engineers, increasing their numbers in the European scientific context
- Offer Open Science and FAIR skills training to policy makers, funders and other relevant stakeholders, like evaluators to promote alignment in evaluation criteria
- Establish policy and coordinate implementation frameworks that measure and reward FAIR and Open Science practises
- Recognise research data and software as official outputs for research for national evaluation of science



### Level 3: Institutional

- Focus on engaging research communities to increase their participation in EOSC through e.g. Rectors' conferences, EOSC reference points
- Incentivise and encourage institutional infrastructures to share best practises on how to channel a broad span of researcher needs into the further development of EOSC
- Offer core Open Science and FAIR skills training to researchers at all levels, taking into account disciplinary differences and different career stages. Such training could be incorporated into doctoral training programmes together with research integrity.
- Embed EOSC information into institutional research programmes to support uptake by research communities
- Foster awareness of local FAIR-enabling practice and the value of long-term preservation of data and metadata amongst research communities
- Support professional development programmes to ensure research support staff have the required Open Science and FAIR skills
- Engage in networks for data stewards, research software engineers and Open Science communities to share models and harmonise best practises
- Address recognition and rewards schemes to provide meaningful career pathways that allow research software engineers and data stewards to progress in academia
- Support the development of an open science policy and associated implementation groups to create an open institutional culture.
- Adjust research review mechanisms and researcher incentives to ensure FAIR research outputs and Open Science are appropriately recognised and rewarded

### Expected outcomes:

- Consolidated information on active EOSC projects is available and regular fora are held to ensure work is aligned and supports the EOSC platform to be embedded
- A well-established stakeholder forum, representative of the various groups and disciplines, ensures EOSC developments meet real-world use cases.
- EOSC has clear engagement from the community with an increasing number of users and service providers
- FAIR, persistent and trusted research outputs and services from a growing range of research communities are available in EOSC
- Validated statistics on national and institutional contributions to EOSC are available, supporting future strategy development
- Researchers and research support teams are equipped with the appropriate Open Science and FAIR skills to transform the way they use research outputs, leading to better-quality and robust research
- Improved availability of highly skilled professionals enabling the practice of FAIR and Open Science through their knowledge of standards, applications and tools
- Researchers are supported by incentives for career advancement that align with open science principles
- Data stewards, Data Scientists, Research Software Engineers and other associated professions are established as new job profiles in the science system
- A higher proportion of research software engineers and other highly-sought staff remain in academia as their skillset is appropriately recognised and rewarded



- Data stewards and researchers can network at a European and national level to access peers for support, consolidating good practises
- The EOSC visual identity is used consistently across projects and the Association
- Open science becomes a national priority and the public national/regional funding agencies use OS principles as a funding condition for research activities.

Coordination across EOSC implementation projects is paramount to ensure developments result in a coherent, operational platform. The EOSC Association should be able to play a key role here. Hence, all new Horizon Europe projects should have a clear obligation written in their consortium agreement to collaborate and engage with the EOSC Association and to contribute to this SRIA. As coordinator of the EOSC Focus project, the Association will try to oversee these activities by running a stakeholder forum, liaising with EOSC projects and hosting project coordination fora. These activities should provide the foundations for the planned liaison necessary to effectively co-create EOSC. The EOSC Focus project will also continue the monitoring function and has activities related to widening EOSC engagement through partnerships with other relevant initiatives. For this to work the EOSC Association should be positioned as the “Data Controller” of the monitoring data.

Widening and embedding EOSC is an activity that applies at all levels. Actions are needed by various stakeholders to help engage research communities and ensure their needs are appropriately addressed. The role of national structures is also key to incentivise research infrastructures to federate into EOSC and support smaller communities on their journey towards FAIR and Open Science practice. While much work has already been done to align EU and national policies with Open Science principles, actions to incentivise the adoption of these practices via support for development of standards and a network of FAIR-enabling trusted repositories will help to roll this out further.

Following on from the implementation of the Rules of Participation, a review should be undertaken to ensure the onboarding procedures are fit for purpose. This should evaluate whether the expected number of national research infrastructures and community services have been federated into EOSC to provide a solid platform from which to build. There should be clear added value for service providers (e.g., integration into EOSC AAI, access to monitoring) and those offering data and code (e.g. increased discovery, ability to combine data from different domains) to embed uptake. Actions at European level and from Member States are needed to continue to incentivize and encourage participation amongst service providers and research communities.

Since a growing number of courses are emerging for FAIR and Open Science, priorities in this phase of implementation focus on aligning curricula to ensure a base-level competency across the European research community. While the European level is responsible for coordinating activity and providing networks to enable data stewards, research software engineers and Open Science communities to engage with peers, much activity focuses on the national and institutional level. It is within these bounds that core competencies can be included in research programmes to ensure a base level of skills. Moreover, national funders and research organisations operate the recognition and rewards programmes and set grant conditions which will be critical to incentivising researchers and support staff and providing appropriate career paths to retain their expertise. Systematic reviews of the new metrics and incentives applied are also needed to ensure they have the desired effect.



## Objective 2

**Enable the definition and adoption of standards, and the development of tools and services, to allow researchers to find, access, reuse and combine results**

### **Priority activities:**

#### Level 1: European

- Maintain and grow entries in the EOSC Interoperability Framework registry, endorsing standards that enable a growing number of communities to connect to EOSC and one another. This registry should act as an exemplar for others on metadata and semantic artefacts.
- Further enhance the EOSC Interoperability Framework to foster composability between services and resources, facilitating the creation of innovative solutions that enable research use cases and foster collaboration between diverse research communities.
- Continue to develop and maintain open interfaces, alignments, crosswalks and APIs to enable interoperability across and inside communities
- Establish clear interoperability guidelines towards the rest of Data Spaces foreseen in the European Data Strategy
- Identify research communities that lack standards and promote their development
- Harmonise standards, tools and services across research disciplines with active participation of research communities
- Define, disseminate and implement common measures of quality to be applied to data, software and other semantic artefacts, so the content within EOSC can be evaluated by users
- Enable fast uptake of emerging technology for EOSC Core to avoid the accumulation of technical debt and thereby ensure seamless operations
- Integrate EOSC services into EOSC-wide AAI systems. Core service integration should be mandatory whilst integration of Exchange services should be strongly encouraged.
- Collaborate with national and thematic competence centres to support adoption of the EOSC Exchange amongst service providers and researchers
- Ensure the quality, long-term availability and usability of the services in the EOSC Exchange by gathering user requirements, disseminating quality information in a transparent way and implementing these by means of detailed specifications, standards adoption, assessment, certifications and monitoring
- Evaluate and as necessary develop trusted environments for managing, preserving and sharing sensitive data
- Develop a global PID resolver and other infrastructure needed to assign, maintain and resolve the broad ecosystem of persistent identifiers used
- Support a network of FAIR-enabling trusted repositories for EOSC, including code repositories, that provide and adopt interoperable tools and aligned services. This network should act as a community platform for shared experience and practices that engages with and uplifts wider data services.

#### Level 2: National

- Encourage and incentivise effective documentation and sharing of all outputs relevant to research, including software and source code
- Provide national Research Infrastructures and Research Performing Organisations with sustainable funding streams related to EOSC, domain and cross-domain initiatives



- Create access to resources from public authorities within EOSC, engaging with national legal and ethical experts to promote appropriate sharing of sensitive data at a European level
- Support and incentivize the development, maintenance and adoption of open standards and APIs to enable resource composability and achieve interoperability across communities including in the private sector.
- Encourage research communities and national RIs to register their interoperability guidelines and where relevant support crosswalks and alignments with others
- Update relevant national data policies (e.g., GDPR, cybersecurity, and similar), by including Open science principles and measures that align with the recognised EOSC standards.
- Incentivise assessment and open certification of data repositories to ensure they enable FAIR outputs, are trustworthy and sustainable, and provide an interoperability layer across the European-wide network

#### Level 3: Institutional

- Raise awareness of infrastructure to support the creation, management and sharing of research data and software, encouraging adoption
- Support research communities to adopt both general and domain-specific standards to increase adoption of FAIR practises and reuse
- Provide guidance and assistance to small repositories to enable sustainability and interoperability through standards compliance, assessment and certification
- Advocate for research communities to support and test the crosswalks and APIs being developed to support interoperability across and inside research disciplines and repositories
- Support the adoption of common quality requirements for FAIR research outputs to be implemented at EOSC level, including through institutional, disciplinary and other data repositories
- Implement tools to plan, track and assess scientific knowledge production (e.g. by connecting science knowledge graphs, Research Information Management Systems, DMPs, metrics) based on open definitions, standards and models

#### Expected outcomes:

- A wide range of disciplinary standards are endorsed and adopted in EOSC and the associated interoperability developments support a growing range of use cases on multi-disciplinary reuse of data and services
- The adoption of metadata standards increases the proportion of FAIR data and range of disciplinary resources that can be discovered via an EOSC search interface
- Infrastructure for sensitive data sharing is provided, via which a comprehensive collection of data from public authorities is made available for research reuse
- A network of trusted data repositories is available and repository managers can identify clear benefits (e.g., efficiency, streamlined operations) of being part of EOSC
- FAIR-enabling trustworthy repositories act as key nodes in the full data lifecycle network of institutional, national and European practice
- Transparency of (meta)data curation and preservation practice across FAIR-enabling data services including trustworthy repositories are mandated and incentivised
- PIDs are used systematically and feed into knowledge graphs to visualise and track research network activity
- Common requirements for the quality of research software and data are agreed and these measures are used to define quality indicators and standards accepted by research communities



- Infrastructures for research software and source code are promoted at national and institutional level, increasing the amount of software and code accessible via EOSC
- Sustainability is secured by national (state/public) institutions and resources, not based mainly on participation in EU funding instruments (e.g., HE programme).

One area of specific interest is extending resources to sensitive data from public authorities. Trusted environments for managing and sharing such data will be offered at the European level and be supported by national level actions to facilitate uptake. A standard set of methods should be developed that can effectively enable data sharing and processing of sensitive content. These should be general enough to be applicable within the same country and in cross-border scenarios. Activities should be cross-articulated with the Data Spaces foreseen in the European Data Strategy to further embed EOSC with parallel digital infrastructure initiatives.

FAIR-enabling repositories for data and code should be brought together in a network to support the development and adoption of standards, tools and practices that promote interoperability and consistency of services. By sharing tools and practises across the network of repositories, less developed services can be supported and advanced. Such activities could facilitate automated metadata uplift, offer multi-lingual knowledge management tools and community curation mechanisms. Certification of repositories is also critical to ensure the services are trustworthy and can be relied upon to maintain the content available via EOSC. Work should also be undertaken to clarify how assessment frameworks for FAIR research outputs and FAIR-enabling trusted repositories work together. This could take into account disciplinary and generic initiatives.

Mechanisms to assess the quality of data, code and other semantic artefacts at the core of EOSC are also needed to ensure that all stakeholders have a clear understanding of the adequacy of the information they are accessing for their scientific application. A data quality framework needs to be defined for EOSC, which includes specifying quality indicators, exploring ways to disseminate certification labels and quality information, providing clear standards to follow and considering potential certification schemes. Guidelines initiated in the EOSC Association Task Forces will be used to steer projects implementing these activities. Institutions will also play a key role to support the adoption of common quality requirements being implemented at EOSC level.

The second phase of implementation activities will focus on growing the research infrastructures and communities connected to EOSC and enhancing the interoperability across them. Collaboration with national and thematic competence centres is envisaged to increase the range of services to add value to the EOSC platform. Inputs from Member States and institutions are needed here to incentivise communities to engage and embed EOSC within a wider user group.

The long-term sustainability of the EOSC infrastructure pivots around the role and mandate of the Association to oversee and steer project implementation. The mandate needs to be formally embedded in project descriptions, grant agreements and procedures clearly defined for efficient execution. Its mandate to take responsibility for project outputs and govern the EOSC infrastructure should also be well-understood and recognised.



## Objective 3

Establish a sustainable and federated infrastructure enabling open sharing of scientific results

### Priority activities:

#### Level 1: European

- Evolve the EOSC governance to effectively coordinate activities and sustain and grow the MVE.
- Widen the resources accessible for EOSC to cover the full spectrum of research outputs (e.g. models, software and workflows) and reliable common services (e.g. AAI, accounting, monitoring)
- Enable researchers across all disciplines and organisations (small and large research infrastructures, and less developed organisations) easy access to the federated infrastructure
- Evolve and further establish the architecture blueprint for the EOSC-Core to enhance the model of the EOSC federated architecture, and as a toolkit to federate research data and enable thematic data spaces in the EOSC Exchange.
- Ensure the EOSC-Core has well-defined technical, governance, and sustainability plans, supported by assessments of critical elements that need high availability and reliability.
- Continue to evolve the EOSC AAI Federation to enable more communities to connect their AARC Blueprint compatible AAIs for seamless access to resources and services
- Enhance the EOSC common digital search functionality to improve discovery, tailored recommendations and reuse of content
- Ensure a continuous security and data protection framework along GDPR other relevant regulatory frameworks with special focus on research and public administrations.
- Support a study on sustainability models for long-term preservation services, which covers the costs of data storage, curation and ensuring long-term usability
- Provide federated long-term access and preservation infrastructure for EOSC, which can be utilised by repositories and research infrastructures. This should include provision of services that support reproducibility of preserved results (e.g. workflows, data processing container images) as well as the data itself.
- Identify resourcing models that enable cross-border access and use of existing national infrastructure capacity
- Trial resourcing models identified in Phase 1 to provide the availability of resources in the MVE, including use and contribution by the wider public and private sector.
- Continue to liaise internationally to develop a global cooperation framework for Open Science infrastructures

#### Level 2: National

- Align research and e-Infrastructure strategies at national level and also between the national and EU levels
- Introduce measures to incentivise RIs and RPOs to adopt appropriate community standards to be federated into EOSC.
- Ensure national investments are directed to national Research Infrastructures that adopt appropriate community and interoperability standards and can be federated into EOSC
- Use national/regional catalogues and pre-production platforms to help providers prepare their resources to be onboarded and scale-out the onboarding process
- Gather information on the financial models implemented by national Research Infrastructures, national research and education networks (NRENs) and communities to

understand any constraints on pooling resources and making services available at a European level

- Continue to harmonise requirements for Data Management Plans (which should address software and other outputs) and encourage the use of machine-actionable DMPs so content can be more easily aggregated and exchanged with other systems

### Level 3: Institutional

- Encourage the adoption and use of the federated infrastructure to be provided by EOSC by sharing examples of how research use cases have been addressed and signposting relevant resources and services when engaging with researchers
- Encourage the use of domain specific and existing Research Infrastructures
- Define and implement procedures to select data, software and other research outputs of long-term value to be preserved via EOSC
- Adopt machine-actionable DMP tools and openly share institutional DMP data where possible in pursuit of establishing best practises and increasing FAIR alignment. DMPs should be interlinked to trusted repositories and other institutional internal systems for managing projects.

### Expected outcomes:

- The EOSC legal structure and governance supports and sustains the MVE and RoP.
- An increasingly broad range of content is available via EOSC and usage statistics demonstrate increased discovery and reuse
- The operation of EOSC Core services (such as AAI) and federated infrastructure model enables a wide range of communities to connect to EOSC
- Consolidation and further enhancement of the EOSC-Core, strengthening the federation aspects and integration
- Long-term preservation infrastructure is offered within EOSC, enabling selected research outputs to be retained and actively preserved for at least 10 years
- Researchers are supported to select data, code and other outputs of long-term value
- Resourcing models to sustain the MVE are trialled and advanced, and a core set of storage and compute resources are offered cross-border
- Data from DMPs is reused to share good practice, increasing adoption of FAIR and making management of research project more effective

Two priority activities within the third objective are supporting long-term preservation and implementing resourcing models for cross-border use of services.

Within the area of long-term preservation, it is recommended to undertake a study on sustainability models. This should investigate the costs and facilities needed, review appropriate file formats to support, consider different models of implementing preservation services and the relative cost/benefit ratio of each, and make recommendations on the minimum retention and preservation periods for different types of data and research software. Decision-making mechanisms to select which data to preserve are also critical and should be elaborated and implemented in conjunction with data services, communities and institutions. Recommendations on long-term preservation services should be implemented by providing effective and cost efficient long-term storage infrastructure at European level to which different repository types can connect.





The EOSC Association Task Force on Financial Sustainability, together with the EOSC Focus project, will explore resourcing models that enable cross-border access and use of services. Financial models to support federated access to existing national capacity are essential to simplify access to resources and avoid fragmentation. This could be combined with in-kind contributions from countries and regions, leveraging results achieved by e-infrastructures which pool capacity and offer easy access across borders. It is recommended to expand the applicability and uptake of the virtual access mechanism across e-infra and research infrastructures based on the experience of the INFRAEOSC-07 projects, as well as exploring other models.

There is an overarching focus under this objective on further developing the federated infrastructure and interoperability as an enabler of Open Science practice. Work should continue to enhance and embed the EOSC Core, enabling a greater range of services and research communities to connect. The common digital search functionality should continue to be enhanced to improve search options and filters. Actions at national and institutional levels will support this by ensuring Research Infrastructures and communities adopt the appropriate standards to be federated into EOSC and make resources discoverable and reusable. Where communities lack standards to contribute to EOSC, initiatives should be supported to promote their development. The role of bodies such as the Research Data Alliance and GO FAIR is important here to link up with international stakeholders and form the European component of the global cooperation frameworks for Open Science. Work to harmonise Data Management Plans and ensure these are machine-actionable is also prioritised as a method to capture and proliferate emerging good practice.