EOSC Future

D2.4a

Project Strategy Plan

The EOSC Future project is co-funded by the European Union Horizon Programme call INFRAEOSC-03-2020, Grant Agreement number 101017536

Version 1.0
December 2021

D2.4a / Project Strategy Plan

Lead by **Technopolis Group Belgium**
Authored by Diego Scardaci (EGI), Gareth O'Neill (TGB), Hilary Hanahoe (RDA), Klaas Wierenga (GÉANT), Ron Dekker (TGB), and Rudolf Dimper (ESRF)
Reviewed by Sarah Jones (GÉANT) & Athanasia Spiliotopoulou (JNP)

Dissemination Level of the Document

Public

Abstract

This is the first version of the Project Strategy Plan for the EOSC Future project. The plan provides a high-level description of the strategy that will guide the project implementation and presents the vision and mission as well as the strategic objectives and approach of the project. The plan focuses on the three overarching tenets of the project to: (1) deliver an operational EOSC Platform (2) integrate and offer resources from the scientific cluster communities (3) actively involve the users in the design and development of the EOSC Platform. A final version of the Project Strategy Plan will be delivered towards the end of the project that will focus on the sustainability and handover of key project outcomes to EOSC stakeholders after the end of EOSC Future.

Version History

Version	Date	Authors/Contributors	Description
V0.1	28/10/2021	Diego Scardaci (EGI), Gareth O'Neill (TGB), Hilary Hanahoe (RDA), Klaas Wierenga (GÉANT), Ron Dekker (TGB), Rudolf Dimper (ESRF)	Table of contents of deliverable finalised
V0.2	18/11/2021	Diego Scardaci (EGI), Gareth O'Neill (TGB), Hilary Hanahoe (RDA), Klaas Wierenga (GÉANT), Ron Dekker (TGB), Rudolf Dimper (ESRF)	Initial draft text of deliverable composed
V0.3	24/11/2021	Carsten Thiel (CESSDA), Gareth O'Neill (TGB), Tiziana Ferrari (EGI)	Draft text of deliverable reviewed and comments incorporated
V0.4	05/12/2021	Gareth O'Neill (TGB), Ron Dekker (TGB), Sarah Jones (GÉANT)	Deliverable quality reviewed and comments incorporated
V0.5	06/12/2021	Gareth O'Neill (TGB), Athanasia Spiliotopoulou (JNP)	Deliverable finalised and circulated to the project consortium
V1.0	09/12/2021	Diego Scardaci (EGI), Gareth O'Neill (TGB), Hilary Hanahoe (RDA), Klaas Wierenga (GÉANT), Mike Chatzopoulos (ATHENA), Ron Dekker (TGB), Rudolf Dimper (ESRF)	Final version of deliverable submitted to the EC

Copyright Notice

This work by Parties of the EOSC Future Consortium is licensed under a Creative Commons Attribution 4.0 International License The EOSC Future project is co-funded by the European Union Horizon Programme call INFRAEOSC-03-2020, Grant Agreement number 101017536.

Table of Contents

Table of Tables

Table of Figures

Glossary

EOSC Future project Glossary is incorporated by reference: https://wiki.eoscfuture.eu/x/JQCK

List of Abbreviations

Acronym	Definition
AAI	Authentication and Authorisation Infrastructure (AAI)
AI	Artificial Intelligence
API	Application Programming Interface
CESSDA	Consortium of European Social Science Data Archives
DIH	Digital Innovation Hub
EAWG	EOSC Architecture Working Group
EC	European Commission
EGI	European Grid Infrastructure Foundation
EIF	EOSC Interoperability Framework
ENVRI	Environmental Research Infrastructures
EOSC	European Open Science Cloud
EOSC-A	EOSC Association
EOSC-SB	EOSC Steering Board
ERIC	European Research Infrastructure Consortium
ESCAPE	European Science Cluster of Astronomy and Particle Physics ESFRI
ESFRI	European Strategy Forum on Research Infrastructures
ESRF	European Synchrotron Radiation Facility
ESWG	EOSC Sustainability Working Group
FAIR	Findable Accessible Interoperable Reusable
GORC	Global Open Research Commons
HPC	High-Performance Computing
IaaS	Infrastructure as a Service
MS/AC	Member States and Associated Countries
MVE	Minimal Viable EOSC
OCRE	Open Cloud for Research Environments
PaNOSC	Photon and Neutron Open Science Cloud
PID	Persistent Identifier
PMB	Project Management Board
PRACE	Partnership for Advanced Computing in Europe
RDA	Research Data Alliance
SDG	Sustainable Development Goal
SME	Small and Medium-Sized Enterprises
SOB	Strategy and Oversight Board
SRIA	Strategic Research and Innovation Agenda
SSHOC	Social Sciences and Humanities Open Cloud

TCB	Technical Coordination Board
TGB	Technopolis Group Belgium
WP	Work Package

1 Executive Summary

The EOSC Future project is a response to the INFRAEOSC-03-2020 call to provide 'integration and consolidation of the existing pan-European access mechanism to public research infrastructures and commercial services through the EOSC Portal'. EOSC Future will realise an operational EOSC Platform that provides access to research products, services, and infrastructures integrated into EOSC to researchers and research communities across Europe by consolidating and scaling up the existing EOSC Portal.

The overarching strategy of EOSC Future can be summarised by three core tenets:

1. realisation of *EOSC-Core* and *EOSC-Exchange* to create an EOSC Platform with interoperable products and services,
2. integration of products and services from the scientific cluster communities in the EOSC Platform,
3. direct involvement of users in the co-design and implementation of the EOSC Platform.

The implementation of EOSC Future will be structured around eight high-level objectives:

1. deliver and operate the *EOSC-Core,*
2. expand *EOSC-Exchange* with resources across disciplines,
3. scale up capabilities and deliver an EOSC Execution Framework,
4. increase European scientific impact with EOSC integrated and interoperable cross-domain scientific resources and collaboration,
5. enable innovation with small and medium-sized enterprises and industry through procurement of commercial services and EOSC Digital Innovation Hub,
6. support and train users and providers of EOSC,
7. reach out and engage with EOSC communities and end users,
8. align implementation with the strategic vision of EOSC.

This report is the first version of the Project Strategy Plan for EOSC Future. The plan is aimed to present an overview of EOSC Future and the high-level strategy that will guide the implementation of the project to EOSC stakeholders and the general public. The plan begins with an introduction to EOSC and the INFRAEOSC-03-2020 call. The plan then goes into detail on the high-level strategy guiding the project before discussing the EOSC Platform, the integration and offer of resources from the science cluster communities and providers for EOSC, and the active engagement of users in the development and co-design of the EOSC Platform. The conclusions provide a short note on the final version of the Project Strategy Plan.

2 Introduction

Scientific research is experiencing an unprecedented digital transition towards open and accessible data. The urgent need for opening up research has been made abruptly clear in the wake of the COVID-19 pandemic. At the core of this transformation lies the general recognition that Open Science contributes to innovation and advancing new scientific discoveries as well as addressing societal challenges more effectively. This global transition goes hand-in-hand with technological developments supporting these advancements. The European Open Science Cloud (EOSC) has been established in this context to support the European Union's drive towards Open Science and tackle the considerable fragmentation, lack of interoperability, and disparate access to quality research data and services for researchers in Europe[1]. EOSC will essentially federate existing research data infrastructures and realise a web of Finadable, Accessible, Ingeroperable, and Reusbale (FAIR) data and related services for science. This web of data and services will make research data interoperable and machine-actionable, allow researchers to find, exploit, and combine linked datasets, provide a basis for artificial intelligence (AI) tools, and lead to new discoveries and new research paradigms.

The realisation of EOSC is envisioned to take place in several stages and has initially been funded via project calls in the framework programme of Horizon 2020. In an initial phase of development from 2017 to 2020, the European Commission (EC) made a financial investment of circa €320 million to build the foundations of EOSC, guided by an implementation roadmap for EOSC[2]. This investment aimed to develop a new pan-European access mechanism to public e-Infrastructures, coordinate related national activities, connect European research infrastructures to EOSC, provide access to publicly funded and commercial services, set up and begin the implementation of the FAIR guiding principles, and start a FAIR-compliant certification scheme for research data infrastructures. To ensure a smooth transition from the first to second stage of EOSC, and involve the wider EOSC community in the preparations, a transition governance was established from 2019-2020 consisting of an EOSC Governance Board, EOSC Executive Board, and EOSC Stakeholder Forum[3]. This resulted in the creation of a Co-Programmed European Partnership for EOSC between the EC and the new EOSC Association (EOSC-A)[4] and a Strategic Research and Innovation Agenda (SRIA) to guide the second implementation phase of EOSC under the Horizon Europe framework programme from 2021-2027[5].

The EOSC Future project is a response to the INFRAEOSC-03-2020 call to provide 'integration and consolidation of the existing pan-European access mechanism to public research infrastructures and commercial services through the EOSC Portal'[6]. This call is foreseen to consolidate and scale up the EOSC Portal[7] and its underlying service platform in order to:

1. strengthen the EOSC Portal so that it continues to provide an increasing portfolio of high-quality standard compliant and interoperable services of proven user interest and scientific relevance from a wide range of national, regional, and institutional public research infrastructures in Europe as well as from commercial service providers in its catalogue,
2. reinforce the role of the marketplace as the access channel to integrated, composable, and reliable services,
3. attract more users, within the research community and beyond, by enhancing the user experience and seamlessly accommodating their needs,
4. ensure its long-term sustainability taking into account all the relevant governance and business frameworks. The project will not only build on the outcomes of key EOSC projects, but will also work closely with the INFRAEOSC-07-2020 projects, which will increase the service offer of the EOSC Portal and set-up a model for interaction between service providers and the EOSC Portal operators[8].

This deliverable is the first version of Project Strategy Plan for EOSC Future. The aim of this plan is to sketch the high-level strategy and principles that will guide the technical implementation of the project as well as coordinate and align activities within the project and with key external stakeholders in the EOSC ecosystem. The plan has been written to inform the wider EOSC community about the project and focuses on three key tenets from the perspective of users. Section 3 presents the vision and mission as well as strategic objectives and approach of the project. The next three sections focus on the key tenets of the project: Section 4 proposes the concept of the EOSC Platform; Section 5 discusses the data and services that will be offered via the EOSC

Platform; Section 6 outlines the plans for user engagement including types of users and how they will be engaged. The plan concludes with a preview of the final version of the deliverable due at the project end.

3 High-Level Strategy

3.1 Vision and Mission

EOSC is envisioned to create a web of research products that will be FAIR (and possibly but not necessarily open) and that will be provisioned with value-added services to fully exploit the products on offer[1]. These products will primarily consist of research publications, data, and code as well as other digital research outputs and training materials. The vision of EOSC Future is to realise an operational EOSC Platform[2] that provides access to the research products, services, and infrastructures integrated into EOSC to researchers and research communities across Europe. This vision is focused on three core tenets:

1. realisation of *EOSC-Core* and *EOSC-Exchange* to create an EOSC Platform with interoperable products and services,
2. integration of products and services from the scientific cluster communities in the EOSC Platform,
3. direct involvement of users[3] in the co-design and implementation of the EOSC Platform.

EOSC Future will enable researchers and research communities to find, access, utilise, and share products from all research disciplines, as well as combine products to reproduce previous research and drive new and potentially cross-disciplinary research.

The creation of the World Wide Web led to the internet becoming the de facto medium for information exchange for both science and society, as well as driving and supporting the interoperability of products and services across technological and domain borders. Infrastructures were subsequently created to utilise the internet for science, including electronic infrastructures to support all communities (such as EGI, EUDAT, GÉANT, OpenAIRE, and PRACE) and research infrastructures to support specific communities (such as the ESFRIs and ERICs). This infrastructural approach has proven successful in supporting the communities but has also resulted in the current culture of communities working in isolation in their disciplinary silos. The mission of EOSC Future is to build on top of this infrastructure by leveraging and strengthening the interoperability of products and services within and across scientific communities. EOSC Future will bring the e-infrastructures and scientific cluster communities (i.e. ENVRI-FAIR, EOSC-Life, ESCAPE, PaNOSC, and SSHOC) together and support communities in making their products and services available across communities, as well as in exploiting other products and services for cross-domain research. This mission aims to realise the three tenets above from the perspective of what will be offered to the users of the platform as shown below in Figure 3.1.

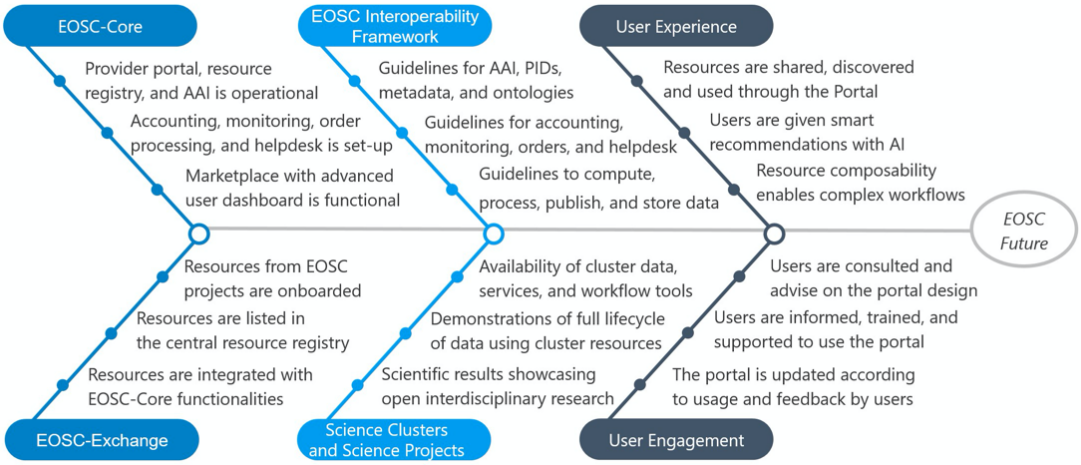

Figure 3.1: EOSC Future Mission to Realise an Operational EOSC Platform

[1] We will use the term 'products' for all digital research objects and 'resources' for both research products and services.
[2] We will henceforth use the more encompassing and actual term 'EOSC Platform' to refer to the 'EOSC Portal'.
[3] We will use the term 'user' broadly for both providers who share their resources and users who exploit available resources.

The *EOSC-Core* and *EOSC-Exchange* will form the technological backbone of the EOSC Platform. The *EOSC-Core* will consist of the minimal services needed to operate the platform. This will include the provider portal, resource registry, and Authentication and Authorisation Infrastructure (AAI), as well as systems for the accounting, monitoring, ordering of resources, and a helpdesk to support queries from users of the platform. The *EOSC-Core* will also include the software for running the marketplace in the *EOSC-Exchange*. Research products and value-added services will be onboarded and integrated with *EOSC-Core* functionalities, listed in a central resource registry, and offered to users via the *EOSC-Exchange*. An intuitive user dashboard will hereby support both the providers of the platform in the onboarding and sharing of their resources as well as the users of the platform in finding and exploiting the available resources on offer.

The EOSC Interoperability Framework (EIF) will act as the glue layer connecting and interoperating the products and services available on the platform. The EIF will provide standards and guidelines for implementing the AAI, Persistent Identifiers (PIDs), metadata schemas, and semantic ontologies. The EIF will also provide guidelines for the accounting, monitoring, and ordering of resources as well as for the user helpdesk. These standards and guidelines will enable the interoperability of resources in the *EOSC-Core* and *EOSC-Exchange*. The initial resources to be made available via the *EOSC-Exchange* will be onboarded from the scientific cluster communities and scientific use cases in EOSC Future. These resources will include research data, workflow tools, and services from the clusters, and will not only demonstrate the benefits of the platform for doing research across the full lifecycle of data using cluster resources, but also showcase the research results from exploiting cluster resources for open and interdisciplinary research. Services from the INFRAEOSC-07-2020 projects, which will offer access to compute, storage, and support for Open Science, will also be onboarded into the platform and offered via the *EOSC-Exchange*.

The user experience is fundamental to the design, functionality, and uptake of the platform. Key users of the platform include researchers who will both share and exploit research resources, but also more broadly resource providers, research and technology enablers, policy makers, trainers, and citizen scientists. These users will be consulted in the design and development of the platform as well as be trained and supported to effectively utilise the platform. A representative user group of EOSC stakeholders will be set up and form a key resource in this engagement. Users will interface with the platform via an intuitive user dashboard[4] that will allow them to share and find resources. Smart recommendations from Artificial Intelligence (AI) embedded in the platform will help them to navigate the web of resources, and combine resources for complex and interdisciplinary workflows. The platform will be updated based on actual user usage and user feedback.

3.2 Objectives and Approach

The vision and mission of EOSC Future will be guided by eight high-level objectives which will collectively contribute to delivering an operational EOSC Platform with research products and value-added services:

1. Objective 1 on 'Deliver and Operate the *EOSC-Core*';
2. Objective 2 on 'Expand *EOSC-Exchange* with Resources across Disciplines';
3. Objective 3 on 'Scale Up Capabilities and Deliver an EOSC Execution Framework';
4. Objective 4 on 'Increase European Scientific Impact with EOSC Integrated and Interoperable Cross-Domain Scientific Resources and Collaboration';
5. Objective 5 on 'Enable Innovation with Small and Medium-Sized Enterprises (SMEs) and Industry through Procurement of Commercial Services and EOSC Digital Innovation Hub';
6. Objective 6 on 'Support and Train Users and Providers of EOSC';
7. Objective 7 on 'Reach Out and Engage with EOSC Communities and End Users';
8. Objective 8 on 'Align Implementation with the Strategic Vision of EOSC'.

[4] We note that there will be a provider dashboard for onboarding resources and a user dashboard for exploiting resources.

These objectives will be coordinated via a unique project governance structure which distinguishes four governing bodies as shown below in Figure 3.2. The General Assembly is the highest decision-making body of the project and consists of all project beneficiaries. The Strategy & Oversight Board (SOB) is responsible for the strategic direction and internal oversight of all project activities and consists of representative beneficiaries predominantly from the e-infrastructure and scientific cluster communities. The SOB oversees the activities of the Project Management Board (PMB) and Technical Coordination Board (TCB) which work closely together. The PMB is responsible for the overall implementation and delivery of the project and consists of project Work Package (WP) leaders. The TCB is responsible for the technological implementation and delivery of the project and consists of high-level technological experts from project beneficiaries.

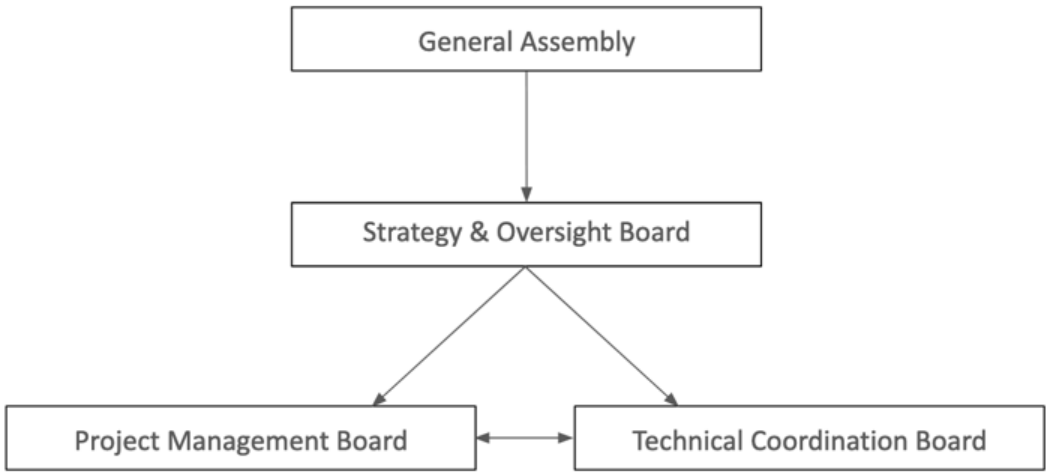

Figure 3.2: EOSC Future Governing Bodies

The SOB leads the project strategy and coordinates internally with the WPs and align externally with strategic stakeholders in the EOSC ecosystem. This coordination and alignment is divided into six thematic pillars which highlight and focus on key aspects of the project and which are led by the members of the SOB:

1. Pillar 1 on **'Policy & Strategy'** defines the strategy and policy of the EOSC platform, ensuring that the project builds upon previous EOSC projects and aligns with the needs of EOSC stakeholders.
2. Pillar 2 on **'Technology & Interoperability'** defines the technological development of the EOSC Platform, ensuring interoperability and alignment with the needs of users and scientific communities.
3. Pillar 3 on **'Excellent Science'** oversees the development of use cases demonstrating how the EOSC Platform supports excellent and cross-disciplinary research to solve scientific and societal challenges.
4. Pillar 4 on **'Co-Development & Procurement'** fosters the growth of the EOSC Platform by connecting resource catalogues and providers and onboarding and procuring public and private sector services.
5. Pillar 5 on **'Skills & Training'** supports training so that users are skilled in using the EOSC Platform, either to onboard and share their resources or to find and exploit other resources for research.
6. Pillar 6 on **'User Engagement'** ensures sharing of information on the EOSC Platform and gathering of feedback from users so that the EOSC Platform is designed and updated to user requirements.

The activities and realisation of deliverables and milestones in the project are structured around 11 WPs:

1. WP1 on 'Project Management';
2. WP2 on 'Project Strategy & EOSC Alignment';
3. WP3 on 'Architecture & Interoperability';
4. WP4 on 'Design & Development of Portal Supply Layer';
5. WP5 on 'Design & Development of Portal Demand Layer';
6. WP6 on 'Integration of Community Services & Products into EOSC';
7. WP7 on 'EOSC Service Delivery & Planning';
8. WP8 on 'Commercial Services';
9. WP9 on 'Training & Skills';
10. WP10 on 'Stakeholder Engagement, Outreach & Marketing';

11. WP11 on 'Ethics Requirements'.

The members of the SOB are responsible for the six pillars whereby each pillar ensures the realisation of related high-level objectives, oversees the activities and internally liaises with related WPs, externally liaises with stakeholders on related topics, and sets and prepares the agendas of SOB meetings. This strategic approach and link between pillars, high-level objectives, and WPs is shown below in Table 3-1.

Table 3-1: Link between EOSC Future Pillars, High-Level Objectives, and Work Packages

Pillar	High-Level Objective	Work Package
Pillar 1 on 'Policy & Strategy'	• Objective 8 on 'Align Implementation with the Strategic Vision of EOSC'	• WP2 on 'Project Strategy & EOSC Alignment'
Pillar 2 on 'Technology & Interoperability'	• Objective 1 on 'Deliver and Operate the ' • Objective 3 on 'Scale Up Capabilities and Deliver an EOSC Execution Framework'	• WP3 on 'Architecture & Interoperability' • WP4 on 'Design & Development of Portal Supply Layer' • WP5 on 'Design & Development of Portal Demand Layer' • WP7 on 'EOSC Service Delivery & Planning'
Pillar 3 on 'Excellent Science'	• Objective 2 on 'Expand with Resources across Disciplines' • Objective 4 on 'Increase European Scientific Impact with EOSC Integrated and Interoperable Cross-Domain Scientific Resources and Collaboration'	• WP6 on 'Integration of Community Services & Products into EOSC'
Pillar 4 on 'Co-Development & Procurement'	• Objective 5 on 'Enable Innovation with Small and Medium-Sized Enterprises (SMEs) and Industry through Procurement of Commercial Services and EOSC Digital Innovation Hub'	• WP8 on 'Commercial Services'
Pillar 5 on 'Skills & Training'	• Objective 6 on 'Support and Train Users and Providers of EOSC'	• WP9 on 'Training & Skills'
Pillar 6 on 'User Engagement'	• Objective 7 on 'Reach Out and Engage with EOSC Communities and End Users'	• WP10 on 'Stakeholder Engagement, Outreach & Marketing'

Sections 4, 5, and 6 below focus on the strategy for realising the three key tenets of an operational EOSC platform, the integration and composability of resources from the scientific cluster communities, and the engagement of users in the design and development of the platform. These sections do not delve into the details of the technical delivery of the project. The technical roadmap, which provides a more detailed planning for the project implementation across key technical areas in three main iterations, is in Appendix A. The roadmap highlights what EOSC Future will technically deliver by project months 6, 18, and 30.

4 The EOSC Platform

4.1 Minimal Viable EOSC

The prevailing definition of EOSC has arisen from the work conducted within the EOSC Sustainability Working Group (ESWG) and the EOSC Architecture Working Group (EAWG)[9]. Via an open consultation process, the ESWG developed the concepts of the *EOSC-Core*, *EOSC-Exchange*, and Minimum Viable EOSC (MVE). These concepts are defined in the Solutions for a Sustainable EOSC, also known as the FAIR Lady, report[10]. On the basis of the FAIR Lady report, and leveraging the outputs of the key EOSC projects (such as EOSCpilot, EOSC-hub, and OpenAIRE Advance), the EAWG developed a high-level diagram depicting the *EOSC-Core*, *EOSC-Exchange*, and MVE in the context of research infrastructures and e-infrastructures (i.e. the EOSC Federation) and the research community at large as shown below in Figure 4.1. The concepts developed by the ESWG and EAWG were properly mapped to the outputs of key EOSC projects that have delivered the current EOSC.

The EAWG has defined the *EOSC-Core*, *EOSC-Exchange*, EOSC Federation, and MVE as:

1. ***EOSC-Core***: The set of enabling services required to operate the EOSC;
2. ***EOSC-Exchange***: The set of federation resources registered to the EOSC by research infrastructures, e-infrastructures, and science clusters to serve the needs of research communities and the widening to the general public and private sector;
3. **EOSC Federation**:The set of resources provided by research infrastructures, e-Infrastructures, and ccience clusters to the research communities.

The **Minimum Viable EOSC** is intended as a dynamic set of EOSC resources:

- The subset of EOSC resources necessary for forming the added value and opportunities considered essential to be provided by the EOSC at a given moment in time, i.e. to allow essential services and research products (e.g. publications, datasets, software) to be discovered, composed, accessed and analysed via the EOSC, which could not be otherwise.
- The subset of *EOSC-Core* components and services required to operate and deliver such resources.

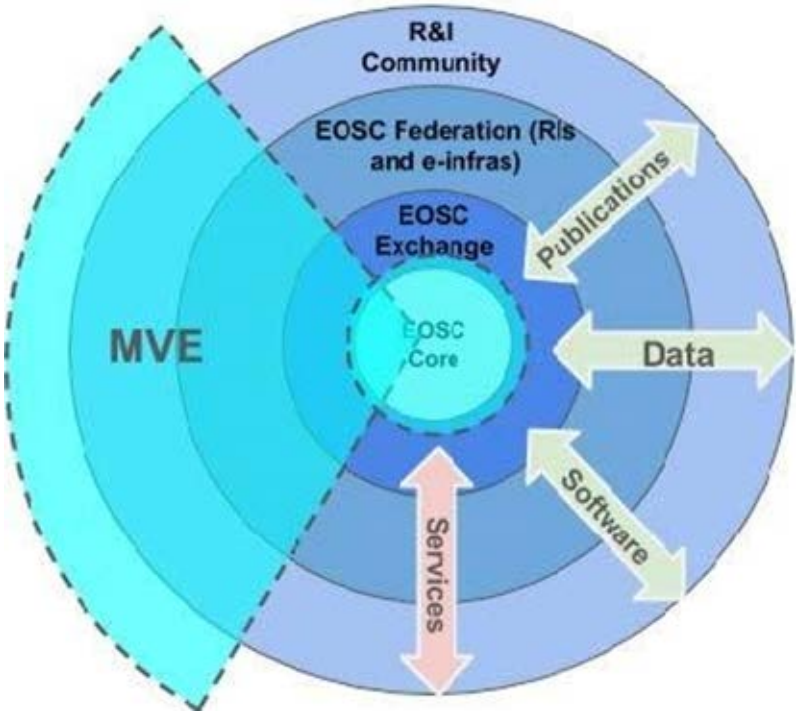

Figure 4.1: High-Level Diagram of EOSC of the EOSC-Core, EOSC-Exchange, EOSC-Federation, and MVE

EOSC Future, building on the solid foundations of the current EOSC which has been collectively assembled and is being operated by a series of key EOSC projects (such as EOSC Enhance, EOSC-hub, and OpenAIRE Advance)

will first bring the MVE into production and will then later expand beyond the MVE to a fuller deployment of the different layers of EOSC as shown above in Figure 4.1.

This will be achieved by enhancing and extending relevant outputs of past EOSC projects and in particular:

- Evolving and expanding the capabilities of the beta *EOSC-Core* handed over by past projects (including EOSC Enhance, EOSC-hub, and OpenAIRE Advance) to meet the expectations of the EOSC Governance, to match needs expressed by researchers and other stakeholders, and to be sustainable.
- Enriching and empowering the *EOSC-Exchange* inherited by past projects (including eInfraCentral, EOSC-hub, and EOSC Enhance) with almost 300 resources already onboarded from major research infrastructures, e-Infrastructures (i.e. EGI, EUDAT, GÉANT, and OpenAIRE), and private companies. This includes scientific tools, horizontal services (e.g. data management tools and container orchestrators), and compute and storage resources. The *EOSC-Exchange* will be enriched increasing the number of available resources covering other research communities and scientific disciplines, while it will be empowered to enable easy composability between its resources and the *EOSC-Core*.
- Scaling up the human support to the users of the platform through the helpdesk and training as well as informing and engaging key stakeholders through the procurement, DIH, and outreach activities.

4.2 EOSC High-Level Architecture

EOSC Future will expand the vision of EAWG and the results and experiences from the previous EOSC projects into a concrete implementation that is designed, implemented, tested, improved, and refined throughout the project lifecycle. This involves not only defining the 'layers' of EOSC, but also the elements within them and how they are inter-connected to form an integrated outcome. A high-level view of the envisioned architecture of the EOSC Platform that will be realised by EOSC Future is shown below in Figure 4.2.

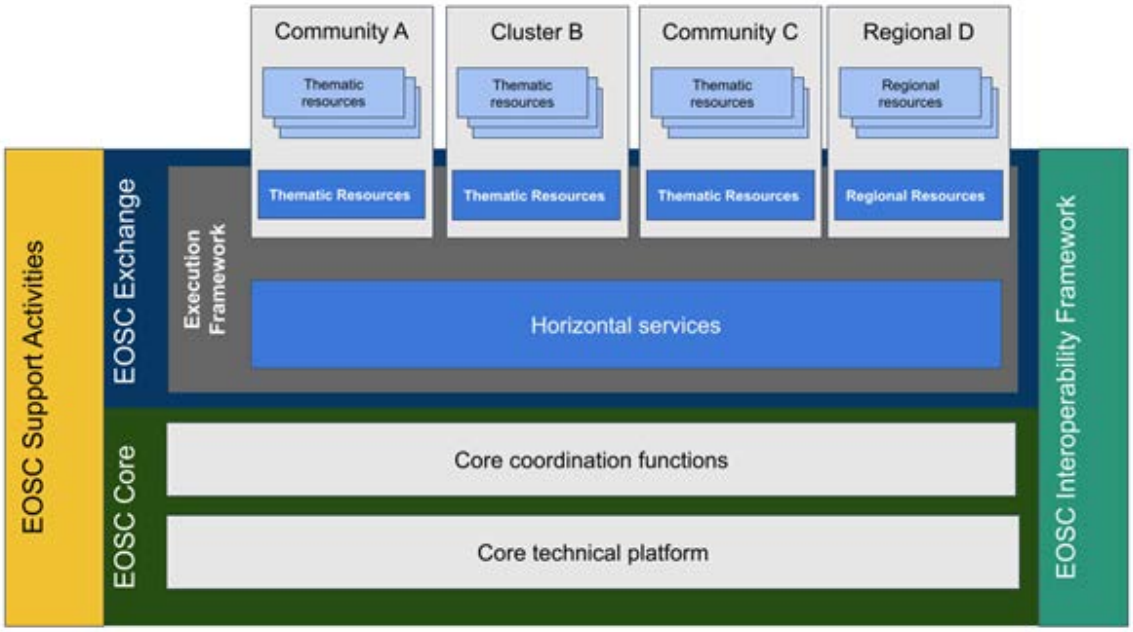

Figure 4.2: High-Level Architecture of the EOSC Platform

The high-level architecture of the EOSC Platform consist of the following components:

1. ***EOSC-Core***: The set of internal services which allows EOSC to operate. This includes a technical platform which facilitates EOSC operations upon which the researcher-facing resources in the *EOSC-Exchange* can rely and integrate with as appropriate. It also includes non-technical coordination functions which operate and facilitate the technical platform (such as the service management system and the onboarding and the security coordination).
2. ***EOSC-Exchange:*** The set of federation services and other resources registered in EOSC by research infrastructures, e-Infrastrucutures, and science clusters to serve the needs of research communities and widening to the public and private sector. Generic services and resources which target

heterogeneous scientific domains and research communities are identified as horizontal services. Resources which target users from a specific discipline, community, and/or regional domain are identified as thematic and/or regional resources. The capability to compose resources across horizontal and thematic and/or regional resources in compliance with the interoperability framework is defined as the Execution Framework. While it is expected that the majority of horizontal services are provided by the e-Infrastructures (e.g. EGI, EUDAT, OpenAIRE, and GÉANT), generic services and resources offered by the science cluster communities will also be offered as a horizontal service.

4.3 System of Systems

The science communities, the e-Infrastructures, and the research infrastructures have over the past 10 years independently built large parts of what is needed for a comprehensive EOSC. Rather than re-inventing the wheel, and accepting that these large user communities will not be willing to abandon their working solutions, EOSC Future will integrate existing research community services and e-Infrastructure offerings into a single virtual platform by adopting a system-of-systems approach. By not building a new system from scratch, the EIF is needed as an additional architectural element to act as a 'glue layer' between these existing systems. The EIF will provide a flexible framework of standards and guidelines to support the interoperability and composability of resources in the *EOSC-Core* and *EOSC-Exchange*. The EIF will act as the glue to connect resources and providers. The EIF is defined as a Reference Architecture Framework which offers the freedom to providers to develop and operate their own provider-specific implementations, while conforming to the EIF guidelines and standards which will facilitate the composition of resources delivered by different providers.

4.4 Implementation

Implementing EOSC in compliance with the above architecture will be achieved via the following approach:

- **Development in EOSC Future is driven by researchers' needs to be successful**. While new offerings may be introduced to researchers, and for which their added value to the researchers' activities may not be known at present, it is imperative that EOSC will serve researchers and help improve the agility, effectiveness, innovation, and impact of research in Europe.
- **EOSC Future will build a system of systems that federates existing digital infrastructures** rather than a new monolithic structure. It must connect, enhance, and foster the greater success of existing providers and research communities, whether from the science clusters, regional projects, national initiatives, or the long tail of science.
- **EOSC Future will reuse not reinvent as far as possible**. A considerable amount of effort has been invested and valuable results have been produced by the existing and previous EOSC-related projects which will form the solid foundations on which EOSC Future will build. This is even dwarfed by the decades of effort within subject-specific communities prior to the initiation of EOSC. There are numerous standards, technologies, approaches, and elements that can be adapted or deployed for EOSC. EOSC Future will seek to reuse these elements where possible.
- **EOSC Future will accommodate the many stakeholder and community members by focusing on the EIF** which will serve as the connecting factor and common language. EOSC Future will be based on previous work done by the EOSC Executive Board Working Groups and in projects such as EOSCpilot, EOSC-hub, and EOSC Enhance. EOSC Future will expand this work by extending the EIF in establishing standards and guidelines to be promoted within EOSC. Via an open process involving research infrastructures, e-Infrastructures, technical experts, and the wider community, standards and guidelines will be established on the basis of consensus. This process will be supported via topical working groups in EOSC Future and in collaboration with relevant EOSC-A Task Forces (e.g. Task Forces on Rules of Participation, Semantic Interoperability, Technical Interoperability)
- **EOSC Future will enhance the *EOSC-Core*** with a suite of internal functions needed to support the operation of EOSC and be offered to resource providers for integration. A strong core will support the expansion of the *EOSC-Exchange* with new resource providers from the science clusters regional projects, and broader European research community, and will facilitate composability of resources.
- **EOSC Future will facilitate the work of researchers by enriching the *EOSC-Exchange*** with a set of key horizontal services that are able to satisfy the needs of multiple research communities and scientific disciplines (e.g. data transfer services, orchestrators, and workflow engines). These

horizontal services will be selected from the portfolio of services of the most advanced European e-Infrastructures and research infrastructures and offered to the wider EOSC audience.

- **EOSC Future will provide researchers with innovative solutions** to enable complex research use cases and foster collaboration between diverse research communities, making resource composability (i.e. the combined or integrated usage of two or more resources) possible. Leveraging the EIF, the EOSC Execution Frameowork will allow orchestrators and/or solutions engineers to build more complete solutions than a single resource and increase the value of EOSC while lowering the barriers for entry.
- Finally, this **overall structure will be strengthened and made more accessible** through the provision of strong technical support, engagement, and training activities to ensure low barriers for entry and realistic support that makes it easier to gain benefit from all the elements described above.

4.5 Functionalities

EOSC Future will develop and implement key components of the EOSC Architecture. The following tables list and describe the main functionalities of the EOSC Architecture that the project will realise. The increasing availability of functionalities over the project's lifetime will ensure that researchers can benefit from the added value of the EOSC Platform in an early stage. Towards the end of the project, a full set of functionalities will be available that support more complex use-cases, such as the capability to discover datasets from disparate scientific disciplines and orchestrate their analysis on adequate (in terms of amount) and appropriate (in terms of type) EOSC compute and storage resources via proper EOSC tools in the EOSC Execution Framework.

Table 4-1: EOSC-Core Platform functionalities

EOSC Architecture Functions	Description
EOSC resource registry and catalogues: services	• The EOSC service catalogues are a part of the overall EOSC resource registry and catalogue ecosystem. • A database of records, for providers, services, and external catalogues, compatible with the EOSC Profiles specifications. • Acessible via the web from the provider portal or via the EOSC provider portal application programming interfaces (APIs) to deposit or update provider and service records. • A content provider to the marketplace, resource registry, or other connected catalogues (e.g. thematic and regional catalogues) that is exposed to researchers, research communities, and third catalogues. • The EOSC service catalogue will incorporate an AI-based recommendation system enabling advanced discovery capabilities.
Data sources in EOSC resource registry and catalogue as a subtype of services within the EOSC profiles	• Data sources are specific EOSC services that provide deposition, preservation, and discovery functionalities for research products. • The EOSC registry data model (I.e. the EOSC profiles) will be extended with additional metadata needed for data sources (for the EOSC registry, provider portal, marketplace, API support). Hence this will likely be combined with 'EOSC resource registry and catalogue: services' in a future version of this list.
EOSC resource registry and catalogues: research products	• The EOSC research product catalogues are a part of the overall EOSC resource registry and catalogue ecosystem. • The EOSC research product catalogue collects and interlinks EOSC research products (such as publications, data, and software) and EOSC services with authors, communities, organisations, services, funders, and projects. • The EOSC research product catalogue supports the onboarding (and validation) of EOSC research product profiles (collected from EOSC data sources) and of EOSC service profiles (collected from the EOSC service registry). It also supports discovery and navigation of EOSC resources as well as statistics.

	• The EOSC research product catalogue will incorporate an AI-based recommendation system. This system will allow advanced discovery capabilities for the EOSC resource catalogues and connected catalogues.
EOSC AAI Federation	• Provides a distributed federated AAI infrastructure which allows for the sharing of login and access to services and data across EOSC.
Connecting *EOSC-Core* components to the EOSC AAI Federation	• Provides shared login capabilities to *EOSC-Core* services and a connector for EOSC-Core platform services to the EOSC AAI Federation.
EOSC order management system	• System to manage orders for services made through the central EOSC catalogue (directly from researcher-facing portal or potentially passed from other catalogues which display services pulled from the central catalogue). • Collect customer/user requests with relevant data, and pass to providers via API, email, or other means. • Support collection of order metrics.
***EOSC-Core* helpdesk**	• Helpdesk platform and support units to support incidents and service requests for the platform services
EOSC monitoring	• The ability to check the status, availability, and reliability of EOSC service resources, both in the EOSC-Core and EOSC-Exchange. • Supports the ability to monitor and observe EOSC resource availability as a measure of resource quality. • Basic monitoring based on service endpoint or web page accessibility. Advanced monitoring based on special probes developed as part of integration with the resources. • Requires configuration management capabilities. • Requires messaging service capabilities
EOSC accounting	• The ability to track and record usage of EOSC resources (i.e. research products and services) both in the EOSC-Core and EOSC-Exchange. • Also EOSC-Core service usage to show the uptake of EOSC-Core services. EOSC-Exchange service accounting based on integration with thematic services and the accounting service to record service usage (e.g. show impact to funders). • EOSC research product usage based on aggregation of usage events collected, according to the COUNTER data usage metrics, from EOSC data sources. • Requires configuration management capabilities. • Requires messaging service capabilities.
EOSC configuration management system	• A shared store of information on which other EOSC-Core services depend (such as monitoring, accounting, and the helpdesk). • Lists resources, resource components, and configuration items which must be controlled (including data about them and changes to them) to effectively deliver EOSC at a federation level.
EOSC collaboration systems	• Internal tools needed to coordinate EOSC-Core and EOSC-Exchange operations. These likely include collaborative information management (e.g. a wiki and document database), task and issue tracking (e.g. a ticketing system), and communication management (e.g. email, chat, mailing list, video conferencing).

Table 4-2: EOSC-Core Coordination functionalities

EOSC Architecture Functions	Description
EOSC security coordination	• Provide community security coordination.

	• Ensure coordination and investigation of security incidents in EOSC-Core platforms and coordinate security incident response activities between EOSC-Core and EOSC-Exchange providers. • Provide guidelines for security risk assessment for providers in the EOSC-Exchange, manage the security baseline for the EOSC AAI proxy and the EOSC AAI Federation, and provide policy guidelines for security, usage, and data supporting secure usage of EOSC resources and exchange of data.
EOSC service management system	• Service management system that allows services and operational roles in delivering EOSC-Core services and supporting external services to be delivered. Includes a set of roles, responsibilities, procedures, policies, and other documentation and tooling to support the management of services.

Table 4-3: EOSC-Exchange and Interoperability Framework functionalities

EOSC Architecture Functions	Description
EOSC helpdesk as a service	• A helpdesk platform where providers can deploy a support unit using their own staff and a helpdesk to support their own services. • Prevents them needing their own technical platform and ensure their help desk is compatible and integrated with the Helpdesk Integration Framework.
EOSC horizontal services	• Services that offer functionalities useful for multiple research communities and scientific disciplines.
EOSC data transfer	• A possible horizontal service to be offered via EOSC-Exchange. • Enable the movement of data files asynchronously between source and destination storage endpoints, including mechanisms to ensure automatic retry in case of failure and for optimisation of performance for large files or large numbers of files.
EOSC research data as a service	• A possible horizontal service to be offered via EOSC-Exchange. • A data access framework combining curated data with supporting crosswalks, tools, services, training and other assets to enable packaging and exploitation of the data. Supports the concept of Data Spaces as supported in Horizon Europe.
EOSC Interoperability Framework management	• The EIF Management is the main instrument to support the integration with the EOSC-Core and realise the resource composability in EOSC. • The resource composability means the combined or integrated usage of two or more resources to provide researchers with added value or innovative solutions. The EIF management takes care to manage, update, circulate and promote the EOSC Interoperability Frameworks, which are created through the projects and task forces to support and enable interoperability within EOSC. • Interoperability includes technical, policy, process and administrative interoperability
EOSC Execution Framework	• The Execution Framework allows the composability of EOSC resources complying to the EOSC Interoperability Framework. This allows orchestrators and/or solutions engineering to build more complete solutions than a single resource and increase the value of EOSC to researchers while lowering barriers to entry. It is built on top of the EOSC-Core services and other resources, and allows them to add value to the offerings in EOSC-Exchange.

Table 4-4: EOSC Supporting Activities functionalities

EOSC Architecture Functions	Description
EOSC web presence	• The EOSC web presence should provide basic information on EOSC, its elements (such as the EOSC-Core, EOSC-Exchange, supporting

	activities and the EIF), its governance, and the opportunities to participate in it.
***EOSC-Core* support**	• Support for use of or integration with EOSC-Core resources, including training and specialist consultancy. • Needed to support researchers to use EOSC-Core elements (e.g. using EOSC AAI to access a thematic service) or by providers (e.g. a thematic provider using EOSC accounting to show usage for Virtual Access repayment).
EOSC Open Science support	• Support for researchers and providers to better embed Open Science practices and services in their work.
EOSC Open Science helpdesk and collaborative tools	• A decentralised and coordinated helpdesk to support service and content providers to share their resources in EOSC and to help researchers in finding support from the most relevant source.
EOSC Open Science metrics	• Access to first-level statistics regarding Open Science, the production of FAIR/open products and FAIRness and openness indicators for research data and research software by organisation, provider, data source, country, community, funder, funding stream, and project.
EOSC DIH	• A structure to co-design and deploy experiments with the commercial sector to benefit from, exploit, or contribute to EOSC. • Support the uptake of EOSC by the commercial sector in multiple ways.
EOSC Observatory	• An interactive dashboard that will facilitate the monitoring of EOSC readiness by MS/AC, indicators for the EOSC Partnership, contributions to the EOSC Partnership and EOSC ecosystem, and national policies on Open Science and EOSC. • The observatory will publicly present results of the monitoring and provide an overview of the implementation of EOSC.
EOSC Knowledge Hub	• The EOSC Knowledge Hub, seamlessly integrated with the EOSC Platform, has two essential components: a training resource catalogue and a learning platform intended to support FAIR sharing and the reuse of training resources in EOSC.

5 Data and Services

5.1 Expanding *EOSC-Exchange* with Resources across Disciplines

In alignment with the FAIR principles and building on the achievements of previous EOSC projects, EOSC Future will bring existing and future tools and services together to create a FAIR data ecosystem beyond research communities and borders. This ecosystem will rely on three key factors to thrive:

1. Add resources to the EOSC platform from the research infrastructures;
2. Add facilities and services from external providers;
3. Increase impact and facilitate use to ensure that researchers will adopt EOSC.

The added value of EOSC is that researchers can search and use data from different scientific domains without the current complexities of searching and finding data in different systems, getting access to multiple data, and entering and understanding different user interfaces. In short, EOSC will eliminate the disciplinary silos of research resources that we are currently faced with. Science clusters and other research infrastructures will onboard their content into EOSC. This will lead to a largely enhanced portal, acting as an entry point for researchers to find and interact with resources and services constituting EOSC. This onboarding of new data (by new providers) should be easy and straightforward, but on the other hand there must be quality checks (e.g. presence of metadata addressing FAIR), and the access conditions for the data must be clear. Moreover, the data should be in a catalogue that is part of the federation of catalogues in EOSC. This will allow researchers to find data sets in research repositories across Europe: a researcher searching on the EOSC Platform will see a comprehensive set of resources from multiple communities and clusters as shown below in Figure 5.1.

Figure 5.1: Science Clusters Onboarding Content into EOSC

Just onboarding data is not enough. It is also important to show the opportunities of combining data from different domains, or to seemlessly connect data with e-Infrastructure facilities. Hence, science clusters will work on specific Science Projects that combine data from different domains, with some projects elaborating the combination of data and e-Infrastructures. The Science Projects will offer test cases focusing on data-driven and mission-oriented research. Although they are case studies, they will require general preparations, like

aligning metadata profiles, combining vocabularies, and setting up user interfaces and secure data spaces, from which other use cases will benefit.

These actions should pave the way for a second added value of EOSC: Composability. Researchers will be able to combine data, construct new data from existing data sets, and add their own data, again not hindered by domain barriers. Besides data, this will require tools and services to be able to exploit the data. The *EOSC-Exchange* will offer such data processing tools and services to analyse, visualise, and combine data from research infrastructures and use e-Infrastructure facilities, such as access to storage, computing, and network services (including services coming from the INFRAEOSC-07-2020 projects). This will also include services from external providers. The *EOSC-Exchange* will offer and recommend services that support data analysis (e.g. based on the data that is selected, a recommendation will be given on which analytic tools can be used).

EOSC Future will accomplish these goals by facilitating the onboarding of new data and service providers, integrating services and connecting with resources, integrating *EOSC-Core* services to data services (to facilitate search and access to data over different domains and providers), and leading by example via science use cases that either combine data over multiple domains or combine data with e-Infrastructure services. The Science Projects will serve as examples of how joint projects can address societal and global challenges and how research infrastructures can align to support Horizon Europe missions within EOSC. The five science clusters ENVRI-FAIR, EOSC-Life, ESCAPE, PaNOSC and SSHOC will run the Science Projects to drive the integration of research infrastructure data and services between the science domains and, as they mature and become ready for production, demonstrate how services from *EOSC-Core* can be incorporated into the routine provisioning of research infrastructure services. The science clusters will contribute thematic services to EOSC, which will be aggregated at the cluster level, reflect the needs, requirements and considerable digital assets and services of the associated research communities, and mobilise communities for widespread use of EOSC resources.

The science clusters will design and implement Science Projects based on the following topics:

- ENVRI-FAIR
  - Climate Change Impact on Biodiversity and Ecosystems in Europe;
  - Dashboard on the State of the Environment,
- EOSC-Life
  - COVID-19 Metadata Findability and Interoperability in EOSC;
  - Imaging Data in EOSC.
- ESCAPE
  - Dark Matter;
  - Extreme Universe and Gravitational Waves,
- PaNOSC
  Tracing Bio-Structures;
  - Dynamics of Biological Processes.
- SSHOC
  - Climate Neutral and Smart Cities;
  - Access Management for Distributed Research Infrastructures.

5.2 Integrating SMEs and Industry and Procuring Commercial Services

Innovation in data-driven research will also come from the outside: new services from 'external' providers will become available and make it easier for researchers to use high-performance digital facilities. One way is to make it easier to use existing (e-Infrastructure) services and to apply them in a new demand-driven environment for the European Data Spaces. This will take place via the INFRAEOSC-07-2020 projects or via SMEs and industry. Hence, this consists of known services such as commercial cloud services, but by focusing these services on the Data Spaces, EOSC Future will create a new demand-driven environment to combine data and services. A second way is to introduce new services (for existing research). This will take place via business pilots

to create collaboration beween EOSC Future partners and SMEs and industy through the Digital Innovation Hub (DIH).

EOSC Future will procure commercial cloud services for EOSC through engagement with providers of services that are required by EOSC users and the development and management of procurement frameworks. Commercial cloud services will be offered via EOSC by providing the procurement function needed to support the delivery of both commodity and tailored services to EOSC users from commercial providers. These services will include Infrastructure as a Service (IaaS), Platform as a Service (PaaS), and Software as a Service (SaaS). EOSC Future will build on previous procurement frameworks for EOSC and will ensure that current and future available commercial frameworks are accessible to EOSC users under favourable terms and conditions. EOSC Future will also procure use cases to achieve data connectivity and data compatibility between EOSC and the Data Spaces. There are nine envisioned Data Spaces: **Industry; Green Deal; Mobility; Health; Finance; Energy; Agriculture; Public Administration; Skills**. These Data Spaces will enable the development of new products and services based on data and include measures for the re-use of public sector information, access to and preservation of scientific information, and guidance on sharing private sector data. The use cases will work towards the broader goal of bringing together data, as a key source of innovation and growth, from different sectors, countries, and disciplines. The use cases will focus on combining and sharing data between businesses, between businesses and science, and between businesses and the public sector.

EOSC Future will further build upon previous efforts to set up a dedicated DIH for EOSC and expand the EOSC DIH as the primary platform for industry usage of EOSC resources and for awareness raising to innovative SMEs and other private enterprises. This will firstly involve the operation and expansion of the EOSC DIH by:

- engaging industry stakeholders to identify new use cases,
- developing strategic partnerships to enhance the services offered via the EOSC DIH,
- serving as an industry engagement channel towards EOSC for other projects and initiatives,
- integrating the EOSC DIH with other established DIHs to support a pan-European network of DIHs, and
- offering support to participating SMEs and DIH partners to identify grant and funding opportunities most relevant to them.

This will secondly involve managing the full lifecycle of onboarded business pilots by:

- analysing requirements of potential business pilots and coordinating EOSC providers and experts to support them,
- developing detailed workplans and exploitation plans with the selected pilots,
- provisioning and enabling access to requested EOSC resources,
- providing first-level support and monitoring progress throughout the pilots,
- supporting the onboarding of services into the *EOSC-Exchange*, and
- identifying business and funding opportunities and defining (pre-)commercial agreements for continued business relationships beyond the pilots.

This will lastly involve the development of a monitoring and evaluation system for the EOSC DIH by defining relevant socioeconomic indicators and metrics to ensure regular assessment of the performance and final evaluation of the EOSC DIH and the business pilots. The EOSC DIH will contribute to the knowledge exchange and alignment within the DIH network and coordinate participation of the EOSC DIH in initiatives such as the DIHnet project DIH Champion Challenge.

6 User Engagement

6.1 Placing Users in the Centre

Aiming for a seamless, easy to use, and practical EOSC Execution Framework, EOSC Future will activate all necessary mechanisms to understand researchers' needs and take appropriate measures to engage them during the design and implementation. The EOSC Future user engagement programme has two main goals:

- Involve users directly and repeatedly in the design and creation of key elements of the architecture (such as standards, APIs, workflows, and accessibility), development, and operations (for scalability, legal, and business).
- Ensure user uptake by promoting the offering and showcasing the benefits of supplying and consuming resources in a stable environment while focusing on the user experience.

A key focus of EOSC Future is to raise awareness among the broader EOSC stakeholder community and engage more users and providers to extend the user base of EOSC:

- **Stakeholder engagement, outreach and marketing activities** will actively reach out to stakeholders via communication and dissemination campaigns as well as engagement events. The RDA community will be targeted to engage in EOSC.
- **Training and skills activities** will develop a dedicated network of trainers as well as training programmes and materials to engage with EOSC Platform users and providers. This will enable them to efficiently provide and exploit resources on the platform and to make effective use of the platform.
- The **use of available data to instigate AI-driven enhancements** will not only attract more users as a result of such innovative services and enhanced user experience, but will lead to a wider adoption of AI and deep-learning techniques on the platform, thus attracting even more users to EOSC.
- **Commercial service activities,** specifically the **EOSC DIH,** will engage industry as co-developers and/or users of EOSC resources via both dedicated engagement and technical piloting activities.
- The project strategy will create **links with high-level EOSC stakeholders** including the EOSC-A, EOSC Partnership, and EU member states and associated countries (MS/AC) as well as a series of essential external stakeholders to support user engagement and awareness.

Given the central focus on users in EOSC Future, all activities will include user-facing goals and objectives, as highlighted in the sections above. The term 'user' identifies all relevant stakeholders of the EOSC Platform and encompasses both those who consume EOSC data or resources and those that contribute to the co-design of EOSC Future solutions by providing resources, sharing requirements, or testing solutions.

The main stakeholders of EOSC Future include the following categories[5]:

- EOSC Consumers:
  - Individual researchers;
  - Research communities;
  - Citizen scientists;
  - Commercial entities (such as private companies interested in using EOSC services).
- EOSC Providers:
  - European e-infrastructures;
  - Research infrastructures;
  - Private sector;
  - Others.
- EOSC Facilitators:
  - Research administrators;
  - Research funders and policy makers;

[5] Complete details on engagement and outreach to users are provided in the public deliverable D10.1 - EOSC Future Stakeholder Engagement and Communication Strategy and Plan (M6).

- Librarians.

6.2 EOSC-Future User Group

An EOSC Future User Group, composed of 300 European researchers and stakeholders as shown below in Figure 6.1, has been created to directly engage and get feedback from users[6]. This group will compliment the internal consortium requirements and feedback gathering mechanisms (including requirements from public and commercial resource providers and clusters) and form the main instrument to collect external requirements and feedback from a diverse range of stakeholders. The user group spans a wide variety of European countries (specifically, 23 MS and 11 AC), includes a diverse range of scientific and disciplinary expertise, and includes more than 75% researchers (both full and part-time).

As a volunteer group, the user group will:

1. complete 4 or 5 short surveys between November 2021 and September 2023,
2. participate in focus groups to provide input on future EOSC Future developments (approximately 2-hour online meetings twice a year between November 2021 and September 2023),
3. contribute to efforts in their scientific community through interaction and feedback from the Science Projects,
4. test EOSC Future solutions before they are rolled out and provide feedback so that the final versions can meet the requirements of the broader user community through two EOSC Future 'testbeds' (i.e. sessions organised between November 2021 and September 2023).

Of note is that upon starting, 59% of applicants are aware or well aware of EOSC, whereas 41% of applicants are not fully aware or do not know about EOSC, and 71% of applicants have never used the EOSC Portal, which underlines the value and uniqueness of this user group to provide feedback to EOSC Future.

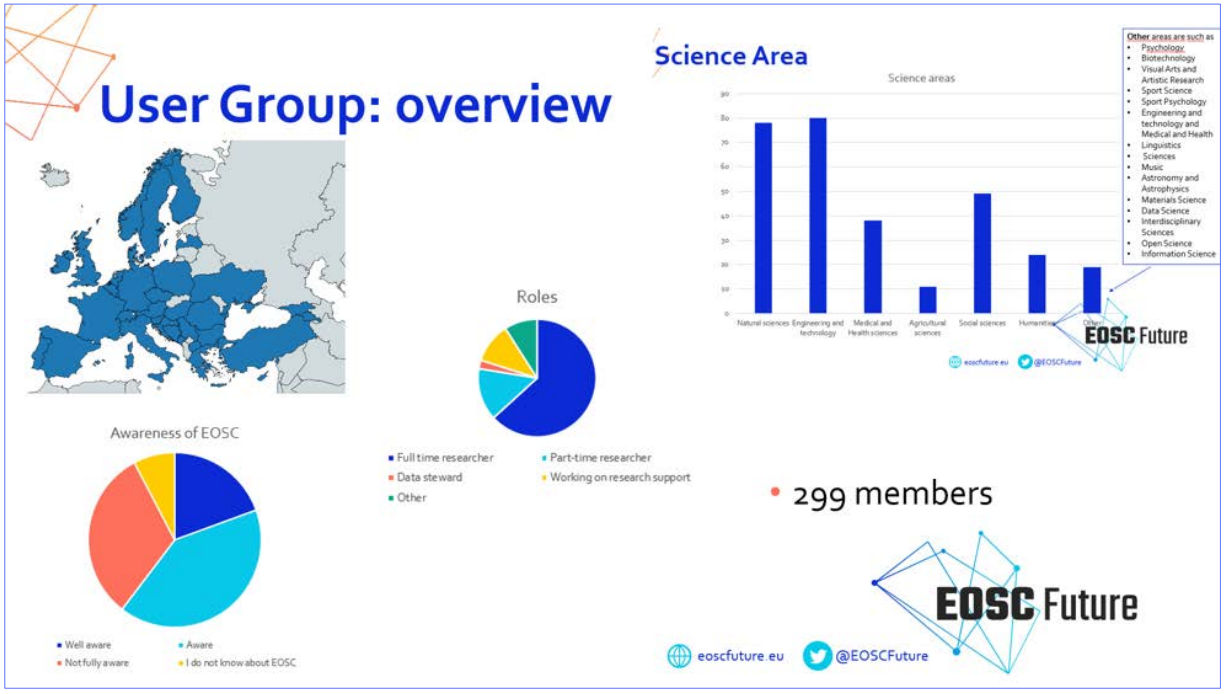

Figure 6.1: EOSC Future User Group Member Overview (as of November 2021)

[6] Complete details in EOSC Future Milestone 33 delivered in M7 (Oct 2021) (MS33 'Co-create EOSC programme launched & user base of co-designers & testers in place').

6.3 User Engagement through Training

EOSC Future will focus specifically on creating training and educational resources and activities for users. The users targeted will be from the same categories as outlined above (i.e. consumers, providers, and facilitators). The functional requirements for the EOSC Future training catalogue and learning platform will be defined through user stories, including learner, trainer, service provider, developer, and funder profiles. The training resource catalogue and learning platform will be released for testing, assessment, and refinement via a number of BETA releases (in months 14, 16, 18, and 20 of the project) and pilots during the project duration, aiming to have the first version of the catalogue in production by month 18 and the beta version of the platform also by month 18. The training activities will consult actively with the user group to identify gaps and (domain specific) areas that require more attention[7].

6.4 User Engagement with External Actors

EOSC Future will engage with other/external initiatives which are part of the wider EOSC ecosystem and whose activities will have significant benefits for EOSC or will have a significant impact on EOSC. This includes liaising with the five INFRAEOSC-07-2020 projects. This also includes close collaboration with the EOSC-A, EOSC Steering Board (EOSC-SB) of MS/AC, and the EOSC Partnership which form the new governace structure of EOSC. This also includes connections with EuroHPC and technical discussions with the hosting sites (including the 3 pre-exascale and 5 peta-scale sites and PRACE) to identify the best ways of enabling accessibility of High-Performance Computing (HPC) resources to EOSC users. EOSC Future will further liaise with GAIA-X, to align the strategic agendas of EOSC and GAIA-X as well as ensuring collaboration and interoperability between the two initiatives, and widening EOSC to the private sector. Finally, EOSC Future will engage and align with the emerging European Data Spaces to ensure interconnectedness between EOSC and the Data Spaces and that EOSC solutions can fulfil the needs of the Data Space communities.

EOSC Future has estabilised a collaboration agreement with the five INFRAEOSC-07-2020 projects:

1. Copernicus - EOSC Analytics Engine (C-SCALE);
2. Data Infrastructure Capacity for EOSC (DICE);
3. EGI Advanced Computing for EOSC (EGI-ACE);
4. Making Open Science More Accessible (OpenAIRE Nexus);
5. Research Lifecycle Management for Earth Science Communities and Copernicus Users in EOSC (RELIANCE).

Three main thematic areas of collaboration have been identified and include technical activities, training, uptake, and engagement. The uptake and engagement thematic area and activity plan will ensure synergies and cooperation with EOSC Future. Activities planned, to leverage on and coordinate the different activities across the six funded projects, include a series of joint awareness raising and promotional actions (including user-focused joint articles, EOSC uptake success stories, and an EOSC provider toolkit).

In terms of joint events, plans include the joint organisation of 'Ask Me Anything' sessions, involving regular (bi-monthly from January 2022) one-hour online sessions to present the EOSC services and widen the EOSC user base. Two EOSC Future provider days will also be organised to encourage resource providers to join EOSC and explain to them how to do that and showcase success stories. An EOSC hackathon-a four-day in-person hackathon on EOSC services-is further planned to encourage the uptake of EOSC resources.

EOSC Future will also bring the data and HPC communities together via EOSC and the science clusters. This effort will help to align strategic agendas and technical solutions, as well as identify the best ways of enabling accessibility of HPC resources to EOSC users, by facilitating the connection between EOSC and the science clusters and the FENIX Infrastructure, a collaboration of HPC centres working on the harmonisation and federation of their offerings of e-Infrastructure services with the goal of supporting a variety of science and engineering communities.

[7] Complete details in D9.1 EOSC Training Catalogue and Platform specification (M6).

6.5 User Engagement with Scientific Communities

EOSC Future has two important mechanisms to engage with the scientific communities for the benefit of users.

The first is with the **ESFRI research infrastructures** which are organised in five thematic science clusters and are directly involved in EOSC Future to deliver:

- cross-domain and cross-infrastructure integration of data, services, and other research outputs for multi-disciplinary science,
- integration with the EOSC Platform and other components of the *EOSC-Core*,
- composability of thematic and horizontal content and resources from e-infrastructures (including the INFRAEOSC-07-2020 projects),
- connect with user communities from five major scientific domains.

The second is through **RDA specific activities and calls** dedicated to widening the awareness and engagement of EOSC in underrepresented communities. This activity will leverage on the research infrastructure and e-Infrastructure experience and best practices matured in EOSC Future, the INFRAEOSC-07-2020 projects, and beyond to support the creation of readiness and awareness of science and research communities that are not included in ESFRI. RDA will leverage on a mapping exercise defined in 2021[11].

The first fifteen communities have been identified:

1. Agriculture, forestry and fisheries;
2. Engineering/wind energy;
3. Engineering/materials sciences;
4. Language and literature/linguistics;
5. Arts/cultural heritage;
6. Health sciences;
7. Bioinformatics;
8. Space sciences/astronomy;
9. Computer and information sciences/librarianship, archival science, and information science;
10. Chemical sciences;
11. Biological sciences/biodiversity;
12. Earth and related environmental sciences;
13. Social sciences/social survey/health ageing retirement;
14. Media and communications/scholarly communications;
15. Public sector information/open government data/Sustainable Development Goals (SDGs).

The RDA Open Call mechanism within EOSC Future to support engagement with a broad spectrum of scientific communities will be used to support user engagement in EOSC, specifically in terms of identifying domain ambassadors, community-specific technical expertise, and RDA disciplinary communities of practice.

6.6 Support for Commercial Services in EOSC

EOSC Future will ensure that researchers have access to a portfolio of commercial services through the EOSC Platform. An initial portfolio of services encompassing Infrastructure as a Service (IaaS) has been established by the Open Cloud for Research Environments (OCRE) project and is currently available via the suppliers listed in the OCRE framework catalogue[12]. This will be extended by EOSC Future based on requirements gathered from users. The EOSC Future DIH activity will focus on business pilots and industry providers. The business pilots conduct joint innovation activities with the DIH for mutual benefit (with a focus on SMEs), while EOSC Future will procure services from industry providers for the benefit of EOSC.

6.7 User Engagement through Collaboration with Global Actors

Being a global community, RDA offers a unique opportunity to EOSC stakeholders to collaborate with leading research-related organisations, research institutes, and funders to extend and sustain their international network and to engage with the best in the world on data interoperability, organisational best practices, and policies in Open Science and research data management. Furthermore, by leveraging on the RDA platform and mechanisms, European stakeholders can access an international forum that facilitates the harmonisation, integration, and/or implementation of standards and methods for research data interoperability, thus providing European research infrastructures with access to tools to deliver research relevant infrastructure.

Plans to ensure international alignment of EOSC as it is developed include the creation of an EOSC Glossary as well as an international research commons glossary which will be synchronised. Further on-going activities include the work conducted by the RDA Global Open Research Commons (GORC) Interest Group and its Global Open Research Commons International Model Working Group[13] as well as the RDA Funders Forum[14] and associated Interest Group[15]. Further activities include coordination with the Data Together Alliance for international activities[16].

7 Conclusions

This Project Strategy Plan has outlined the strategic view of EOSC Future and proposed approach to realise the three key tenets of the project; to **build an operational EOSC Platform, integrate and offer resources from the science cluster communities and relevant providers for EOSC,** and **actively engage users in the design and development of the EOSC Platform**. This first version of the plan will strategically guide the technical implementation of the project so that the EOSC Platform is fully operational by the project end in month 30 in September 2023. The updated version of the plan will be delivered towards the end of the project and will focus on the sustainability and handover of project outcomes to key EOSC stakeholders after the project. It will also focus on the project outcomes that need to be sustained, recommendations to sustain these outcomes in the long-term, and identifying which outcomes will be handed over to key stakeholders.

8 Appendix A – EOSC Future Technical Roadmap

Area	By M6	By M18	By M30
User Experience - Resource Sharing and Discovery	• Researchers can see and reach all thematic and regional portals from the EOSC Portal. • Researchers can see services from the thematic clusters through EOSC. • A researcher from PaNOSC can seamlessly use compute and/or storage resources provided by the e-Infrastructures to analyse data from the PaNOSC Research Infrastructure using the PaNOSC (UmbrellaID) identity and without having to re-register across infrastructures. • A researcher from a new community not involved in EOSC can get a 'virtual tour' of available resources.	• A researcher searching on EOSC Portal sees a comprehensive set of resources from multiple communities and clusters. • Usage statistics for datasets (views, downloads) will be collected and made available. • Researchers using resources through EOSC will have common elements such as AAI, support, monitoring, and accounting. • A researcher from PaNOSC can access an ESCAPE resource with the PaNOSC (UmbrellaID) identity. • Researchers using a thematic portal see resources (services,data) pulled in from the central EOSC registry. • A richer set of horizontal services to support science is offered to researchers.	• A researcher can do the full lifecycle of data processing, storage, analysis, and publishing supported by resources available and transparently integrated through EOSC. • EOSC allows research communities to build cross-disciplinary portals. • Researchers can gauge the quality and suitability of resources based on usage statistics and feedback from other services.
User Experience - Resource Allocation	• A researcher can request resources offered by INFRAEOSC-07 projects through the EOSC Marketplace.	• A researcher requesting resources through the EOSC Marketplace portal can request access to EC-funded resources.	• Requesting resources through EOSC includes the possibility to access commercial or centrally funded resources. • Researchers can request HPC resources (e.g from EuroHPC) directly through the EOSC Marketplace.
User Experience - Resource Composability	• Researchers can see example cases of complex workflows using multiple resource providers. • Researchers can access and use EOSC computing and storage resources. • Researchers can see a rich range of horizontal resources and request access to them.	• Researchers can orchestrate data analysis on computing resources provided by multiple e-Infrastructure resource providers and transfer back the output to his/her storage system.	• Researchers can compare and select resources based on how easy they are to compose and connect to. • Communities can offer their users fully integrated (end-to-end) workflows and a number of these are available for various research topics.

Area	By M6	By M18	By M30
EOSC-Exchange	• Resources onboarded through prior projects (e.g. EOSC-hub, EOSC Enhance, eInfraCentral) remain available. • All horizontal resources from 07 projects are onboarded. • A selection of resources from prior projects and 07 projects are integrated with functions from *EOSC-Core*. • Initial cluster services have been onboarded and are listed in the Resource registry in the EOSC portal and accessible through the EOSC Marketplace. • Horizontal services (publishing workflows, data transfer, data packaging, container deployment) are demonstrated through 07 service instances. • Resource requests are passed to the provider.	• First release of EOSC File Transfer Service as first service arising from the clusters to become a horizontal service. • Majority of cluster resources (services, data) are available through the central resource registry and EOSC Marketplace. • All main 07 horizontal services are integrated with *EOSC-Core* functionalities. • Numerous other horizontal services (publishing workflows, data transfer, data packaging, VM/container deployment and orchestration) are integrated with *EOSC-Core* functionalities. • *EOSC-Exchange* offers AI/ML-enabled suggestion functions for using the portal(s). • *EOSC-Exchange* highlights which services can be easily integrated or composed. • The onboarding process is extended to include validation of data sources to align to community (FAIR) metadata guidelines. • The onboarding process for resources is extended to not only list them but also to include optional integration steps in the same workflow. • Resource requests can lead to automatic provisioning of resources.	• Production release of EOSC File Transfer Service. • First releases of other horizontal services arising from clusters and communities. • Ability to create thematic execution environments/VREs based on integration of compliant thematic, horizontal, and core resources. • The onboarding process allows for automated/self-service integration with some *EOSC-Core* functionalities. • The onboarding process for resources is extended to include more optional integration steps in the same workflow. • Resource requests integrated with procurement as well as provisioning functions. • *EOSC-Exchange* includes numerous services from communities other than those represented by the clusters.

Area	By M6	By M18	By M30
EOSC-Core	**EOSC Marketplace** • Shows all data collected from provider and resource profiles. Allows for scoring and rating. **Provider Portal & Resource Registry** • Allows direct onboarding (web + API) but as transfer of provider and resource records from other registries and vice-versa. **EOSC AAI** • EOSC AAI Federation operational. **Monitoring & Accounting** • Allows checking of services based on availability of their web pages/endpoints. Can support better integration via specific metrics. Accounting supports usage tracking to support virtual access reimbursement by the EC. **Helpdesk** • Helpdesk covers core services. Customisation based on new/updated technology and connected to the Service Management System. **Order Processing** • Order management is set up and includes integration of providers. Already used for services inherited from EOSC-hub and others that previously enabled it. **EOSC Portal Metrics Dashboard** • The dashboard provides statistics about services and requests of access and is fully integrated with the EOSC Portal.	**EOSC Marketplace** • Connected to monitoring to be able to enrich feedback with availability data. Includes AI/ML-based suggestion engine interfaces to automatically pull in resources from other catalogues in the EOSC central registry. **Provider Portal & Resource Registry** • Increased automatic validation tools and automatic flagging of resources which are likely to require review. **EOSC AAI** • e-Infrastructure SP-proxies and cluster community AAIs fully integrated to EOSC AAI Federation. Community AAIs can integrate. **Monitoring & Accounting** • Automated/self-service integration of monitoring probes and metrics offered to providers. Monitoring can track availability and reliability and accounting of usage based on parameters in provider and resource profiles (location, sector, organisation type). **Helpdesk** • Helpdesk covers core services and can redirect tickets to providers who have their own helpdesk. Helpdesk can be tested by providers who wish to use it. **Order Processing** • Order management can deploy and provision resources from selected providers, including from 07 projects. **EOSC Portal Metrics Dashboard** • The dashboard is enriched with additional information on EOSC resource providers and on activities of researchers in EOSC.	**EOSC Marketplace** • Describe supported interfaces, standards, workflow languages, and metadata supported by EOSC resources. An advanced user dashboard is available. **Provider Portal & Resource Registry** • Inbuilt or integrated management/workflow engine to support management of applications, review of records, auditing and quality control as well as automatic provider communication. **EOSC AAI** • Community AAIs seamless integration with EOSC AAI federation through self-service onboarding. **Monitoring & Accounting** • Automated monitoring includes automated thresholds, raising issues, or alarms in the Service Management System based on results. Accounting is related to capacity data offered by providers. **Helpdesk** • Helpdesk-as-a-service available as optional add-on during onboarding. Integrated with central helpdesk functions. **Order Processing** • Order processing for thematic services can bundle orders for horizontal or basic services needed to deliver the thematic service and include deployment and provisioning. **EOSC Portal Metrics Dashboard** • The dashboard provides statistics related to the combined/integrated usage of EOSC resources.

Area	By M6	By M18	By M30
Clusters and Science Projects	**SP Deployment** • Preparation phase of SPs: cluster SPs started integration and adaptation of cluster/Research Infrastructure tools and services into the broader EOSC framework, including integration with AAI, data sources moved into EOSC (FAIR) data stores, and catalogued, software development and exchange platforms available, workflow deployment mechanisms adapted. **Clusters' Input to EOSC Horizontal Services** • List of candidate services from clusters which can be generalised to be EOSC-wide offerings, with several already being developed to be EOSC-ified.	**SP Deployment** • Initial deployment of SPs done: Most of the SPs have fully operational workflows using integrated EOSC services and tools, make use of EOSC Interoperability Framework, provide feedback from the science communities to the service deployment and operation, and can use resources made available through the clusters/Research Infrastructures. **Clusters' Input to EOSC Horizontal Services** • EOSC software catalogue/repository populated with identified cluster-provided services and tools.	**SP Deployment** • Full scale operation of SPs to the point where many have full scientific analyses ready or close to publication as full demonstrations of open cross-disciplinary science. Demonstrations of the full lifecycle of data processing, storage, analysis and publishing supported by resources available and transparently integrated through EOSC. Workflows deployed across cluster/Research Infrastructure resources and where appropriate on commercial cloud and/or European HPC resources. **Clusters' Input to EOSC Horizontal Services** • General availability of Research Infrastructure-originated horizontal services (as appropriate) visible through EOSC Portal and catalogues.

Area	By M6	By M18	By M30
EOSC Interoperability Framework	**Resource Description Framework** • Provider and resource description framework v3.2 including data sources. **Identifiers** • Initial overview of available PID frameworks and guidelines for selecting PID types. **AAI** • Initial technical guidelines for Research Infrastructures and e-Infrastructures to connect AARC-compliant AAI Proxies to the EOSC Federated AAI. EOSC AAI Federation guidelines accepted by the cluster communities and e-Infrastructures. **Metadata and Ontologies** • Initial guidelines for metadata discovery and exchange on the basis of existing generic guidelines (e.g. OpenAIRE, DataCite, EUDAT, DCAT). **Accounting** • Initial guidelines for reporting accounting metrics for virtual access by INFRAEOSC-07 projects. **Monitoring** • Initial guidelines for monitoring service URLs registered in the EOSC Catalogue. **Order management** • Initial guidelines for managing orders specifying interfaces to forward orders to providers. **Helpdesk** • Initial guidelines for handling user requests for services registered in the EOSC Catalogue. **Data Platforms for Processing** • Initial guidelines for data ingesting and movement for processing in hybrid cloud environment. **Data Publishing and Open Data**	**Resource Description Framework** • Provider and resource description framework v3.5 including research products and interoperability guidelines/best practices for horizontal services in *EOSC-Exchange*. **Identifiers** • Initial guidelines for new PID types (e.g. instruments, services, software, organisations) and standards to connect PID frameworks to PID Graphs. Draft guidelines for PID service providers for minimum Kernel Type Information **AAI** • Initial technical guidelines to connect IdP and AAI proxies from public and private sector service providers to the EOSC Federated AAI should become technical guidelines for cross-infrastructure credential delegation and verification for supporting multi-step agent-driven workflows. **Metadata and Ontologies** • Initial guidelines for communities to publish community-specific metadata and ontologies in EOSC. **Accounting** • Initial interoperability framework for service providers for automated reporting of accounting metrics for VA. **Monitoring** • Initial interoperability framework for service providers for monitoring service availability. **Order Management** • Initial interoperability framework for service providers for automatic dispatching of the orders for services registered in the EOSC Catalogue. **Helpdesk** • Initial interoperability framework for service providers for automatic dispatching and handling of user requests for services	**Resource Description Framework** • Provider and resource description stable release v4.0 incorporating new features requested by the user and provider communities. **Identifiers** • Draft interoperability framework for a PID meta-resolver and guidelines for PID service providers for minimum Kernel Type Information. **AAI** • Technical interoperability guidelines for supporting cross-sector access to the EOSC Federated AAI. **Metadata and Ontologies** • Guidelines for minimum metadata to support the discovery, metadata exchange, and cross-walks of research products across communities. **Accounting** • Extended interoperability framework for service providers for automated reporting of accounting and usage metrics. **Monitoring** • Extended interoperability framework for service providers for monitoring service availability of services registered in the EOSC Catalogue. **Order Management** • Extended interoperability framework for service providers for automatic dispatching of orders for services registered in the EOSC Catalogue. **Helpdesk** • Extended interoperability framework for service providers for automatic dispatching and handling of user requests for services registered in the EOSC Catalogue. **Data Platforms for Processing** • EOSC-endorsed guidelines for data ingesting and movement for processing in hybrid cloud environment adopted by

Area	By M6	By M18	By M30
	• Initial guidelines for data repository. **Cloud Compute Containerisation and Orchestration** • Initial guidelines for VM/container management and orchestration. **HTC-HPC Compute** • Initial guidelines for HPC/HTC clusters on demand and multi-tenant containerised job submission. **Machine Learning** • Initial Interoperability guidelines for Machine Learning and review of existing interoperability guidelines (e.g. from EOSC-hub, OpenAIRE Advance, FAIRsFAIR).	registered in the EOSC Catalogue. **Data Platforms for Processing** • Guidelines for data ingesting and movement for processing in hybrid cloud environment improved according to user communities' feedback. **Data Publishing and Open Data** • Guidelines for data repository improved according to user communities' feedback. **Cloud Compute Containerisation and Orchestration** • Guidelines for VM/container management and orchestration improved according to user communities' feedback. **HTC-HPC Compute** • Guidelines for HPC/HTC clusters on demand and multi-tenant containerised job submission improved according to user communities' feedback. **Machine Learning** • Guidelines for Machine Learning/Deep Learning data analytics services improved according to user communities' feedback.	one or more horizontal services. **Data Publishing and Open Data** • EOSC-endorsed guidelines for data repository adopted by one or more horizontal services. **Cloud Compute Containerisation and Orchestration** • EOSC-endorsed guidelines for VM/container management and orchestration adopted by one or more horizontal services. **HTC-HPC Compute** • EOSC-endorsed guidelines for HPC/HTC clusters on demand and multi-tenant containerised job submission adopted by one or more horizontal services. **Machine Learning** • EOSC-endorsed interoperability guidelines for Machine Learning.

References

[1] Webpage on European Open Science Cloud (EOSC) hosted by the European Commission. Link: [https://ec.europa.eu/info/research-and-innovation/strategy/strategy-2020-2024/our-digital-future/open-science/european-open-science-cloud-eosc_en]. Accessed 24 November 2021.

[2] European Commission. Implementation Roadmap for the European Open Science Cloud (2018) Link: [https://www.eoscsecretariat.eu/sites/default/files/files/swd_-_implementation_roadmap_for_the_eosc.pdf]. Accessed 24 November 2021.

[3] Webpage on EOSC Governance 2019-2020 hosted by the EOSCsecretariat.eu project. Link: [https://www.eoscsecretariat.eu/eosc-governance]. Accessed 24 November 2021.

[4] European Commission and EOSC Association. Memorandum of Understanding for the Co-Programmed European Partnership for the European Open Science Cloud (EOSC) (2018) Link: [https://www.eosc.eu/sites/default/files/EOSC_Memorandum_30_July_2021.pdf]. Accessed 24 November 2021.

[5] EOSC Association. Strategic Research and Innovation Agenda (SRIA) of the European Open Science Cloud (EOSC) (2021). Link: [https://www.eosc.eu/sites/default/files/EOSC-SRIA-V1.0_15Feb2021.pdf] Accessed 24 November 2021.

[6] Webpage of EOSC Future on CORDIS hosted by the European Commission. Link: [https://cordis.europa.eu/project/id/101017536]. Accessed 24 November 2021.

[7] Website of the EOSC Portal hosted by the EOSC Future project. Link: [https://eosc-portal.eu]. Accessed 24 November 2021.

[8] Webpage of the INFRAEOSC-07-2020 call on CORDIS hosted by the European Commission. Link: [https://cordis.europa.eu/programme/id/H2020_INFRAEOSC-07-2020]. Accessed 24 November 2021.

[9] Webpage of the EOSC Executive Board Working Groups hosted by the EOSCsecretariat.eu project. Link: [https://www.eoscsecretariat.eu/eosc-working-groups]. Accessed 24 November 2021.

[10] EOSC Executive Board. Solutions for a Sustainable EOSC (2020) Link: [https://op.europa.eu/en/publication-detail/-/publication/581d82a4-2ed6-11eb-b27b-01aa75ed71a1]. Accessed 24 November 2021.

[11] Webpage of the RD4EOSC project hosted by the Research Data Alliance. Link: [https://www.rd-alliance.org/rda4eosc-researchcommunities]. Accessed 24 November 2021.

[12] Webpage of the OCRE cloud catalogue hosted by the OCRE project. Link: [https://www.ocre-project.eu/services/cloud-suppliers]. Accessed 24 November 2021.

[13] Webpage of the Global Open Research Commons Interest Group hosted by the Research Data Alliance. Link: [https://www.rd-alliance.org/groups/global-open-research-commons-ig]. Accessed 24 November 2021.

[14] Webpage of the RDA Funders Forum hosted by the Research Data Alliance. Link: [https://www.rd-alliance.org/about-rda/organisational-bodies/rda-funders-forum]. Accessed 24 November 2021.

[15] Webpage of the Research Funders and Stakeholders on Open Research and Data Management Policies and Practices Interest Group. Link: [https://www.rd-alliance.org/groups/research-funders-and-stakeholders-open-research-and-data-management-policies-and-practices-ig]. Accessed 24 November 2021.

[16] Data Together Alliance. Proposal for a Data Together Alliance. Link: [https://www.rd-alliance.org/sites/default/files/attachment/Data%20Together%20Final%20Version%20March%202020%5B5%5D.pdf]. Accessed 24 November 2021.