



EOOSC-hub

Briefing Paper - EOOSC Federating Core Governance and Sustainability

Status	Final
Version	Public
Delivery Date	18/07/2019
Dissemination Level	Public

About this paper

This Briefing Paper contains an initial proposal of the EOOSC Federating Core, illustrating a possible approach to its composition and relating it to functional and non-functional requirements emerging from EOOSC use cases. Initial proposals for its governance and sustainability are also put forward, provided as input towards the ongoing implementation of the European Open Science Cloud.

Delivery slip

From	<p>Editor: Dale Robertson</p> <p>Contributors and Reviewers: Magchiel Bijsterbosch, Tiziana Ferrari, Luciano Gaido, Damien Lecarpentier</p>
-------------	---

COPYRIGHT NOTICE



This work by Parties of the EOSC-hub Consortium is licensed under a Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>). The EOSC-hub project is co-funded by the European Union Horizon 2020 programme under grant number 777536.

GLOSSARY

<https://wiki.eosc-hub.eu/display/EOSC/EOSC-hub+Glossary>

<i>Terminology/Acronym</i>	<i>Definition</i>
EGI Federation	A federation of computing and storage resource providers united by a mission to support research and development. The federation is governed by the participants represented in the EGI Council and coordinated by the EGI Foundation.
EOSC	The European Open Science Cloud promoted by the European Commission to provide all researchers, innovators, companies and citizens with seamless access to an open-by-default, efficient and cross-disciplinary environment for storing, accessing, reusing data, tools, publications and any EOSC Resource for research, innovation and educational purposes ¹ .
EOSC Executive Board	Body of representatives from the research and e-infrastructures communities, appointed by the European Commission ²

¹ <https://eoscpilot.eu/eosc-glossary#overlay-context=eosc-glossary>

² <https://www.eoscsecretariat.eu/eosc-governance/eosc-executive-board>

EOSC Governance	Overall Governance Structure for EOSC, comprising EOSC Governance Board, EOSC Executive Board and Stakeholder Forum (latter not yet specified)
EOSC Governance Board	Also “EOSC board”: institutional group gathering the member states and the Commission to ensure effective supervision of the implementation
EOSC-hub	Project creating the integration and management system of the future European Open Science Cloud
EUDAT CDI	European e-infrastructure of integrated data services and resources to support research
FAIR	Guiding principles to make data Findable, Accessible, Interoperable, and Reusable
GÉANT	Pan-European research and education network that interconnects Europe’s National Research and Education Networks (NRENs)
Horizon 2020	The European Union Framework Programme for Research and Innovation
INDIGO-DataCloud	Project developing a data/computing platform targeted at scientific communities, deployable on multiple hardware, and provisioned over hybrid (private or public) e-infrastructures
OpenAIRE-Advance	Project supporting Open Access/Open Data mandates in Europe
PRACE	Partnership for Advanced Computing in Europe

Contents

Executive Summary	5
1. Introduction	7
2. EOSC Federating Core	9
2.1 Introduction	9
2.1.1 Recommendations for the Federating Core	10
2.1.2 Recommendations for the EOSC Service Portfolio	11
2.2 EOSC Federating Core	13
2.2.1 The Hub Portfolio	13
2.2.2 Compliance framework	14
2.2.3 Shared Resources	15
2.3 EOSC in the Research Cycle	17
3. Federating Core Services	19
3.1 Hub Portfolio	19
3.2 Compliance Framework	22
4. Sustainability, Governance and Costing	24
4.1 Levels of EOSC Value-Add	24
4.2 EOSC Access	24
4.3 Sustainability and Governance	25
4.4 Indicative Costing	26
4.5 Access and Compensation	28
5. Concluding Remarks	30
Appendix I. EOSC-hub Contribution to the Hub Portfolio and Compliance Framework	Error!
Bookmark not defined.	
Hub Portfolio	Error! Bookmark not defined.
Compliance Framework	Error! Bookmark not defined.

Executive Summary

Building on the recommendations of the EOSCpilot project, the High-Level Expert Group reports, and the EOSC-hub use cases, this document proposes an approach to the definition and the provisioning of the EOSC Federating Core, which will be based on existing capabilities of e-Infrastructures and research infrastructures at national and European level in an open and participatory manner.

It is proposed that the Federating Core delivers three capabilities:

- (1) Technical services and support activities that implement the federating tier (Hub portfolio) to provide integration³, i.e. the activities and tools to provide coordinated access and management of resources (services and scientific products) provided by multiple suppliers (internal and external to EOSC organisation). EOSC resources are expected to be delivered at national and European level, together with the support and expertise necessary to address complex digital needs of the EOSC user communities. The Hub portfolio delivers the EOSC “**federating tier**”.
- (2) Shared resources, which include the scientific outputs (data, applications, software, pipelines etc.) and the storage and compute hosting platforms needed to deposit, share and process these outputs. The shared resources realise the EOSC “**resource tier**”.
- (3) The Compliance Framework, providing the EOSC “**regulatory tier**” (including the Rules of Participation, the Service Management System and related policies), that define the policies and processes for the demand side and the supply side to engage with EOSC.

We propose the Federating Core to be complemented by the EOSC Service Portfolio which will provide additional added-value services (common and thematic) which exploit the Federating Core and are discoverable, selected, customised and instantiated through the EOSC Portal, to address the needs of specific user communities. The composition of the Federating Core and the EOSC Service Portfolio will be driven by EOSC-defined Rules of Participation, technical and policy requirements that will define the EOSC conformance requirements for providers.

Together, the EOSC Federating Core and EOSC Service portfolio will enable the management of the full research data lifecycle.

Following the HLEG recommendations, we propose that the entire EOSC Federating Core is entrusted to the governance of the EOSC. Funds should be made available to sustain the costs of federating existing digital infrastructures – leveraging existing investments, and expanding their capacity. Funding should come from member states, and multistate entities (like the EC, international research organisations etc) to support Excellence-Driven and Wide Access models. Market-driven access will be sustained by paying customers, like research projects and commercial entities.

- We propose the Hub Portfolio and Compliance Framework elements of the Federating Core to be part of the “Minimum Viable Ecosystem” entrusted to EOSC and accessible through the

³ Integration could include the definition of policies, interoperability, guidelines, standard protocols, reference open APIs, minimum metadata, and tools for demand and supply-side aggregation like a portal, and resource catalogues

Wide Access mode to all suppliers participating in the EOSC federation. These two components, central to the EOSC, are necessary to build the integration and interoperability framework across the services and scientific products that are currently provided through multiple different policies and access channels. Therefore, we propose that the components or solutions which provide the Hub Portfolio and Compliance Framework services are governed by the EOSC. Their sustainability would therefore be synonymous with that of the EOSC itself.

- We propose that Shared Resources provide multiple access modes, namely: Wide access, Excellence-driven, and Market-driven, depending on the user segment served and economic characteristics of the resource. Long-term funding to the Shared Resources will maximise the exploitation of shared scientific products and create economies of scale across the member states. In particular, we propose:
 - Wide access: for user communities requesting on-demand access to limited capacity, and scientific application platform managers contributing to EOSC open access by adding valued data information services, and requiring long-term provisioning of compute, storage and data hosting capabilities from EOSC.
 - Excellence-driven access: for research collaborations and the long tail of science requesting a sizeable amount of resources for a limited period of time.
 - Market-driven: for long-term research collaborations, projects and commercial exploitation.

A **partial ballpark indication** of the cost of activities and services of the Hub Portfolio and the Compliance Framework, based on those relevant activities which are within the scope of the EOSC-hub project, is approximately 74 FTEs per annum, equivalent to around 7.6 million Euros per annum. This does not include any costs in the resources tier and will have to be updated once a complete set of capabilities has been defined and agreed by the EOSC community.

Finally, we propose that the services of the EOSC portfolio are independently owned by their respective providers. These services are made discoverable and accessible through the EOSC portal, using EOSC as an “invisible coordinator”.

This briefing paper will be the subject of public consultation. We propose this briefing paper to initiate a community activity aiming at discussing and evolving the Federating Core concept, that should be ultimately reflected in a community position paper.

1. Introduction

European Open Science Cloud (EOSC) refers to the initiative promoted by the European Commission *to provide all researchers, innovators, companies and citizens with seamless access to an open-by-default, efficient and cross-disciplinary environment for storing, accessing, reusing data, tools, publications and any EOSC Resource for research, innovation and educational purposes*⁴.

The EOSC landscape is evolving rapidly, with contributions, all aimed at building a strong and sustainable offer, including the March 2018 EC EOSC Implementation Roadmap⁵, the November 2018 final report of the second EOSC High Level Expert Group⁶, and the EOSCpilot project which ended in April 2019. The EOSC governance was announced in November 2018 and the EOSC Portal was also launched. With the support of the EOSC Secretariat and the EC, these developments are being further built upon by the EOSC implementation projects and related initiatives like RDA. The EOSC governance is currently setting up five Working Groups, on Architecture, Landscape, Sustainability, FAIR and Rules of Participation. There have also been significant developments in the EuroHPC Joint Undertaking⁷, an important part of the European Data Infrastructure which is a parallel and complementing initiative announced alongside the EOSC as one of the three pillars of the European Cloud Initiative⁸. The EOSC-hub project is developing its activities against this backdrop.

The funding bodies both at national and EU level play a crucial role in the development of a sustainable EOSC. To this end, the EOSC business proposition needs to be defined, taking into account EOSC use cases, and existing recommendations leveraging the efforts of the EOSC High-Level Expert Group and the EOSCpilot project. The EOSC offer needs to be underpinned with a funding framework that provides adequate support for long-term planning and service provision. The European Union is in the process of finalising the next Multi-annual Financial Framework (MFF) which will include funding and accompanying instruments to finance research and innovation activities including the Horizon Europe Programme.

The challenges of creating a sustainable EOSC are such that they can be tackled only by an overarching collaboration of EOSC stakeholders. The definitions of the EOSC funding model, organisational model and other governance-related topics need to be widely discussed and solved at the level of Member States and the EC.

Task 2.3 of the EOSC-hub project is responsible for developing and delivering recommendations for a Governance and Sustainability Roadmap, leveraging the input of major national and European e-Infrastructures and Research Infrastructures participating in the consortium. EOSC-hub also

⁴ <https://eosc-portal.eu/glossary>

⁵ SWD(2018)83 final:

https://ec.europa.eu/research/openscience/pdf/swd_2018_83_f1_staff_working_paper_en.pdf

⁶ Prompting an EOSC in Practice:

<https://publications.europa.eu/en/web/eu-law-and-publications/publication-detail/-/publication/5253a1af-ee10-11e8-b690-01aa75ed71a1>

⁷ EuroHPC Work Plan 2019: <https://eurohpc-ju.europa.eu/documents/EuroHPC-Work-Plan-2019.pdf>

⁸ <https://ec.europa.eu/digital-single-market/en/news/communication-european-cloud-initiative-building-competitive-data-and-knowledge-economy-europe> The three pillars are the EOSC, the European Data Infrastructure and Widening Access and Building Trust.

collaborates with OpenAIRE-Advance to produce a proposal on aligned roadmaps for service positioning and sustainability within EOSC (Joint Activity Milestone 3.4 in the EOSC-hub/OpenAIRE-Advance Collaboration Agreement). As a step towards this, and as input to the European Commission to assist with programming of work in the next MFF period, this Briefing Paper presents initial proposals for the governance and sustainability of the EOSC Federating Core.

The proposals are put forward for further discussion, and feedback on the contents of this paper is welcomed. An open consultation will be launched to collect feedback from stakeholders. We propose this briefing paper to initiate a community activity aiming at discussing and evolving the Federating Core concept, that should be ultimately reflected in a community position paper.

The structure of the briefing paper is as follows. The composition of the EOSC Federating Core is proposed in Chapter 2, reflecting the recommendations from the EOSCpilot, the High-Level Expert Group and the EOSC-hub use cases. The functionality of the proposed services is described in Chapter 3. Chapter 4 estimates funding requirements and provides recommendations for the initial sustainability of EOSC. Chapter 5 concludes the paper with information about consultation on these initial proposals and next steps.

2. EOSC Federating Core

2.1 Introduction

The purpose of this Chapter and the following one is to provide proposals for an initial definition and organisation of the capabilities (functional and non-functional) the Federating Core of EOSC should deliver. The methodology adopted has been based on the analysis of recommendations stemming from the HLEG final report, the EOSCpilot recommendations, and the requirements and use cases emerging from research communities from within and outside the EOSC-hub consortium. The communities from the EOSC-hub consortium are represented by nine thematic service providers and eight competence centres^{9,10}. These communities include various ESFRI projects and landmarks, international communities of practice and international research projects¹¹. They bring input from a broad range of scientific disciplines. External user communities include evidence-based policy makers and the EuroGEOSS European initiative, commercial entities engaged through the EOSC Digital Innovation Hub, Virtual Research Communities, the Open Research Cloud Alliance initiative and one HPC Centre of Excellence.

According to the EOSCpilot, *“the essential findings and requirements from the experience of the Science Demonstrators include various recommendations. They start with the requisite of a close and continuous collaboration between EOSC and the active research communities. The suggestions follow with the need for high-level services for research end-users and easy access to large-scale resources and services. In addition to this, it is critical that domain-aware experts are available as support when using EOSC services. Last but not least, addressing the FAIR data principles depends on a community by community basis”*¹². Similarly, the HLEG report “Final report and recommendations of the Commission 2nd High Level Expert Group on the European Open Science Cloud (EOSC)”, noted that *“An MVE [Minimum Viable Ecosystem] will emerge and thrive only if some basic technical, political and human resources conditions are met. From a technical perspective, interoperable services and open data must be guaranteed. On a human resources level, a coordinated effort has to be made to put the right incentives in place for all of those involved (researchers, software developers and infrastructure managers, research managers) to design, contribute to and exploit the system. Policies are needed for*

⁹ Disciplines and initiatives covered are:

- physical sciences - astronomy (LOFAR), fusion (ITER), high-energy physics (CMS and VIRGO), space science (EISCAT-3D)
- earth science (EO pillar, GEO) – climate research (ENES), seismology (ORFEUS)
- biological sciences – marine and freshwater biology (IFREMER-led thematic service), biodiversity conversation (LifeWatch), ecology (ICOS)
- humanities – language and literature (CLARIN), arts (DARIAH)
- engineering – environmental engineering (sea vessels, LNEC services), civil engineering/others (disaster mitigation services)
- medical and health sciences – biological sciences (ELIXIR), structural biology (WeNMR)

¹⁰ For further information see D7.2 First Report on Thematic Service Architecture and Software Integration <https://www.eosc-hub.eu/deliverable/d72-first-report-thematic-service-architecture-and-software-integration>

¹¹ ESFRI landmarks represented (number of infrastructures): energy (1), environment (5), health (6), physical sciences and engineering (11), social sciences and humanities (5), other (1)

¹² <https://www.eosc-pilot.eu/content/d44-consolidated-science-demonstrator-evaluation-report>

the technical and human-oriented conditions to emerge.” In addition to this, the report highlights that: “The researchers’ job is based on data and on computational resources. They need to produce or find data relevant to the inquiry, find an appropriate service or hosting for their own data and services, do the necessary transformations, run the analysis, publish the results and make data available to others. Flexible ways to access and share data and direct access to fast networks to do so are at the top of the agenda for researchers. One of the main problems that the EOSC needs to solve is the fact that researchers in Europe still have insufficient access to e-infrastructures.”

2.1.1 Recommendations for the Federating Core

It is proposed that the Federating Core delivers three capabilities:

- (1) Services that implement the federating tier (Hub portfolio) to provide integration¹³, i.e. the activities and tools to provide coordinated access and management of resources (services and scientific products) provided by multiple suppliers (internal and external to EOSC organisation). EOSC resources are expected to be delivered at national and European level, together with the support and expertise necessary to address complex digital needs of the EOSC user communities. The Hub portfolio delivers the EOSC “**federating tier**”.
- (2) Shared resources, which include the scientific outputs (data, applications, software, pipelines etc.) and the storage and compute hosting platforms needed to deposit, share and process these outputs. The shared resources realise the EOSC “**resource tier**”.
- (3) The Compliance Framework, providing the EOSC “**regulatory tier**” (including the Rules of Participation, the Service Management System and related policies), that define the policies and processes for the demand side and the supply side to engage with EOSC.

The Federating Core is complemented by the EOSC Service Portfolio which provides additional added-value services (common and thematic) which exploit the Federating Core and are discoverable through the EOSC Portal.

The composition of the Federating Core and the EOSC Service Portfolio will be driven by EOSC-defined Rules of Participation, technical and policy requirements that will define the EOSC conformance requirements for providers.

As a consequence of this, it is recommended that the overall EOSC Resource Portfolio includes two broad categories of resources: the EOSC Federating Core and the EOSC Service Portfolio, as illustrated in Figure 1.

¹³ Integration could include the definition of policies, interoperability, guidelines, standard protocols, reference open APIs, minimum metadata, and tools for demand and supply-side aggregation like a portal, and resource catalogues

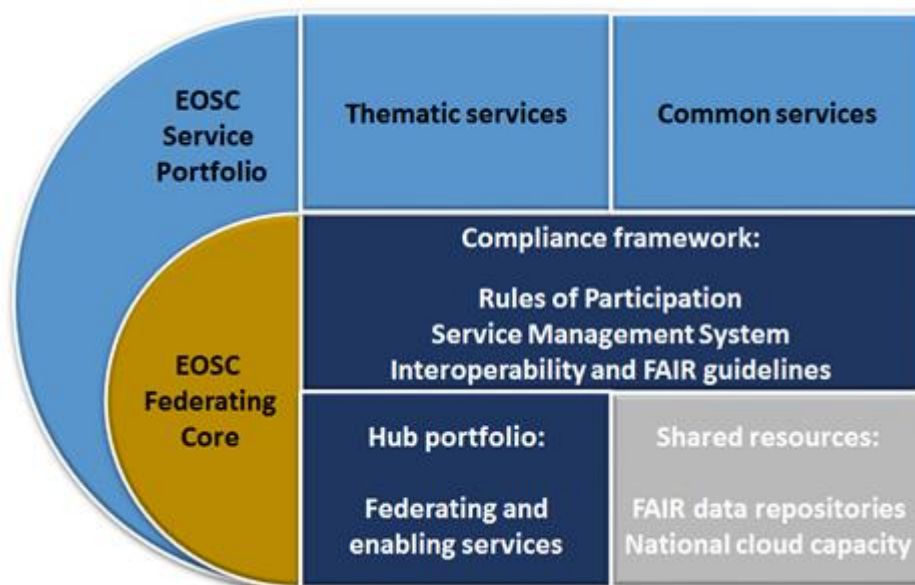


Fig. 1 - The EOSC Resources, organised into two portfolios: the EOSC Federating Core (yellow) and the EOSC Service Portfolio (light blue).

The first element, the **EOSC Federating Core**, is a fundamental asset that EOSC-hub proposes to the EOSC. It includes a number of parts which were already described in the EOSC Implementation Roadmap (SWD/2018/83) and have been expanded and instantiated by EOSC-hub, resulting in a more complete picture of the elements that are needed to allow the research targeted services to operate. These elements are technical, human, policy and resource related by nature and they must be maintained over the long term.

The Federating Core comprises the resources entrusted to EOSC, namely

- The Hub portfolio contributing the EOSC federating tier, i.e. providing access- and federation-enabling services (Section 2.2.1). Examples of Hub Portfolio services are federated AAI and the EOSC Portal
- The Compliance framework providing the EOSC regulatory framework, providing the policies and processes required to operate the Federating Core (Section 2.2.2).
- Shared resources contributing to the EOSC resource tier. This category comprises resources (services and scientific products) of pan-European relevance which are developed by a given discipline but used more broadly by external user communities and additional disciplines (Section 2.2.3).

The composition of the Federating Core and the EOSC Service Portfolio will be driven by EOSC-defined Rules of Participation, technical and policy requirements that will define the EOSC conformance requirements for providers.

2.1.2 Recommendations for the EOSC Service Portfolio

The second element, the **EOSC Service Portfolio**, provides added-value resources (both thematic and common) making use of the EOSC Federating Core, and providing complementary capabilities to the EOSC users.

-
- The **Thematic services** of the EOSC Service Portfolio are community-specific capabilities including research core data, data products, scientific software, and pipelines. They are provided by science centres, Research Infrastructures and other sources, supported by national or international infrastructures. Examples of thematic services are: data resources and software tools to access, study and compare the data; data brokering services tailored to the needs of specific scientific communities; climate change analytics; shared virtual environments for finding and using Earth Observation data; molecular simulation tools for drug discovery, etc.
 - The **Common services** (sometimes referred to as a “horizontal” catalogue) provide generic capabilities usable by any science discipline, each supporting aspects of the data lifecycle from creation to processing, analysis, preservation, access and reuse. Examples of services belonging to this category are multi-disciplinary services for data discovery, processing, workflow management and orchestration, data management, curation and preservation, and for data access, deposition and sharing.

The services of the EOSC Service Portfolio are owned and provided by a variety of public and private service providers. EOSC supports their discoverability by promoting them to international user groups to promote multidisciplinary science and broader exploitation. Services of the EOSC portfolio are made discoverable through the EOSC Portal, supported to conform to EOSC guidelines and best practices, but are externally accessed and funded according to the policies and business models of choice of the respective providers. EOSC services are provided and managed under the responsibility of their respective providers.

We recommend the policies contained in the Compliance framework of EOSC to apply equally to both the EOSC Federating Core and the EOSC Service Portfolio.

The HLEG final report presented an overall user analysis that includes four User / Provider roles: End User, Service Developer, Research Funding Organisation and Core Infrastructure Provider.

- The End User is expected to interact with EOSC to: register for authentication and authorisation purposes, describe data, discover services and data, transform data, run analysis, store results and publish. ⇒ The federating tier and the resource tier of EOSC provide the environment for this.
- The Service Developer is expected to create new services based on Shared Resources, and publish services and provide consulting services to relevant end users. ⇒ Service Developers contribute to the development of the EOSC Service portfolio and can benefit from a resource tier of scientific outputs that would not be otherwise available in Europe.
- The EOSC Core Infrastructure Providers are the selected research infrastructures and e-Infrastructures responsible for supplying the resource and federating tiers to EOSC.

2.2 EOSC Federating Core

2.2.1 The Hub Portfolio

We envisage **the Hub Portfolio** (also noted as EOSC Platform in some EC documents) to include the portfolio of key access- and federation-enabling services needed to operate EOSC – both technical and human. Such services would be, for example, authentication and authorisation infrastructure, accounting, the marketplace, and other elements which enable the federation, access, ordering, delivery, reporting and management of research-facing services and the shared resources, as well as training, technical and policy support to users in EOSC-relevant areas. We anticipate that additional capabilities will be provided by other EOSC projects and initiatives, and we will work with them to contribute to an updated Hub Portfolio description.

As well as operating as part of the Federating Core, the Hub Portfolio services can also be offered to EOSC Service Portfolio providers who do not already have these components, or do not wish to operate them themselves. For instance, the helpdesk can be provided to them as a component to integrate into their service.

The Hub portfolio not only includes services for federation, but also the human network of trainers and supporters that is necessary to provide advice and concretely enable access to complex digital environments. As suggested by the final HLEG report, *“the generation of expertise in deploying and running advanced services to support frontier research creates know-how in the resource centres. Often this type of service is only available in research infrastructures as prototype, long before it becomes commercially viable or profitable, if ever. When fed back into industry, in the form of trained people, it is this know-how that deliver the added value necessary for economic growth.”* This recommendation was reinforced by the finding of the EOSCpilot project. Scientific Demonstrators highlighted the *“need for close and continuous collaboration between EOSC and the active research communities”*, the *“clear benefit of robust bi-directional communication structures”* and the *“importance of IT and domain aware experts to support researchers”*¹⁴

An example to illustrate how the EOSC Hub Portfolio may be of benefit to a researcher is provided in the box below.

The EPOS Use case for Federated AAI¹⁵. A seismologist requests to access the services of the EPOS-ORFEUS. He (or she) is redirected to the community Authentication Service (relying on an EOSC AAI solution in the background) where he can log in at his home institution or create a local account if needed. He receives a token. Depending on the profile he might be authorised to use the services. Profiles include information about the groups he belongs to (e.g. read permission of particular restricted data). After this, the researcher (authenticated and authorised) wants to perform an analysis on a dataset previously selected and staged. He logs in to the Jupyter environment close to the staged datasets. The researcher selects and launches a kernel containing his preferred seismological libraries. When the corresponding Jupyter notebook is up and running the datasets

¹⁴ D4.4: Consolidated Science Demonstrator evaluation report (<https://eosc-pilot.eu/content/d44-consolidated-science-demonstrator-evaluation-report>)

¹⁵ Source and credits: the EPOS-ORFEUS Competence Centre, EOSC-hub project

are available in a local directory and he can perform his analysis. He might choose to pause his work and save it for later. Finally, he can download results on his PC, move them to his personal cloud storage folder or make them available on a local folder.

2.2.2 Compliance framework

Another element of the Federating Core is the **Compliance framework**, which represents the policies and processes required to operate the Federating Core. The main vehicles are the Rules of Participation, the EOSC-hub Service Management System, and the Interoperability Guidelines for thematic and common services, but other policies may be included on a case-by-case basis, in particular when they concern collaboration with other EOSC implementation projects. Beyond this, EOSC-hub is trying to develop shared policies and protocols for programmatic exchange of information with other EOSC stakeholders, on topics such as ordering, reporting and accounting. Allowing a flow of information between stakeholders supports interoperability and a better user experience.

The Rules of Participation is a key element of the Compliance framework. It sets out the policies to be adhered to in order to provide thematic or common services through the EOSC Service Portfolio. The intention is not to limit access to services to the research market, but rather to ensure that the services provided can be understood, ordered, assessed and reported on in a coherent way. The Rules of Participation set out and oversee the on-boarding process for thematic and common services, and manage exceptions or new developments that require changes to policies and procedures.

The other main element, the EOSC-hub Service Management System helps to ensure that the services in the EOSC Service Portfolio and in the Federating Core are managed in a professional, predictable, measured and optimised way. The Service Management System is organised according to the FitSM standard that is compatible with, but more lightweight than, the traditional ITIL framework and ISO/IEC 20000 standard. The EOSC-hub Service Management System comprises 14 processes which will be described in more detail in *D2.6 First Service roadmap, service portfolio and service catalogue*, along with further information on their applicability to the Service Portfolio. We anticipate that the Service Management System scope will be focussed on the Federating Core services entrusted to the EOSC, and to apply to EOSC Portfolio Services only in a lightweight manner, for example to validate them and enable their registration in the EOSC Portal.

Finally, several Interoperability Guidelines are under development and will assist smooth service on-boarding and management. They include, for example, technical guidelines for common and thematic services, federated service management guidelines, and service description metadata, the latter being part of a joint effort with the eInfraCentral and OpenAIRE-Advance projects.

An illustration of the value the Compliance Framework could provide to researchers is provided in the box below.

The LOFAR Use Case: Ingestion of FAIR data¹⁶. As a radio-astronomer, the scientist wants to enter science-grade data products in a science data repository that supports the FAIR principles to ensure

¹⁶ Source and credits: the radio astronomy Competence Centre of EOSC-hub

long-term data preservation and attribution of effort. This will further improve sharing of data with colleagues and access to data from other science domains. It should be possible to access data in the science data repository using direct links to individual data objects via an anonymously accessible public URL such that other services, e.g. those provided by the Virtual Observatory, can be built to provide access to the data.

2.2.3 Shared Resources

The third element suggested to be incorporated into the Federating Core portfolio is the **Shared resources**. While the Hub portfolio enables access to services and supports the federation, and the EOSC Service Portfolio onboards thematic and common services, the Shared resources would provide services such as those for data management and processing for EOSC exploitation, hosted by a generic resource tier of storage and computing and scientific products. The following use cases provide examples of how the long tail of science and major Research Infrastructures may benefit from them.

Data

Sharing of EU Copernicus data for research¹⁷. The massive streams of high-resolution Earth Observation (EO) data derived from the EU Copernicus sensors have established Europe as the predominant spatial data provider for use in global environmental monitoring applications. These data are made available under a full, free and open license with an unprecedented frequency and spatial extent. In principle, the availability of these new data sources should lead to their integration in a wide range of science and monitoring applications spanning regional to continental scales. The latter is, in fact, happening but primarily outside of Europe, in particular facilitated by large US IT companies. This leads to the unfortunate situation where large EO user communities, in particular in the science domain, need to rely on non-European platform suppliers for the necessary Big Data Analytics that is required to scale high volume use of the data streams. Whereas expertise in EO analysis solutions and use is world class and widespread in Europe, to date there are no solutions that provide core cloud services coupled to an online long-term data archive of Sentinel¹⁸. Hence, there is a need in Europe for:

1. Easily accessible European computing environments that allow for rapid scaling and sharing of (Sentinel) data among a large community of users.
2. A European platform, similar to Google Earth Engine, which can make large scale storage accessible through sophisticated indexing and caching solutions with an advanced application programming interface (API).

Applications

Scientific applications as a service for structural biology. The structural biology community of

¹⁷ Source and credits: The European Open Science Cloud for Earth Observation Science Community, G. Lemoine, JRC (<https://indico.egi.eu/indico/event/4431/session/31/contribution/67/material/slides/0.pdf>)

¹⁸ <https://scihub.copernicus.eu/>

practice WeNMR¹⁹ needs user-friendly tools for its users, hiding the complexity of computing resources and to ensure sufficient resources to operate them. WeNMR has a long history of using high throughput computing resources under EGI. Maintaining the quality of WeNMR services and support, together with continuously adapting and improving them (e.g. to make use of new compute models or facilitate their use through the implementation of single-sign-on mechanisms) is a constant challenge. The EOSC is being set up to be Europe's virtual environment for all researchers to store, manage, analyse and re-use data for research, innovation and educational purposes.

Computing

Scaling up of in-house ICT infrastructures with EOSC for the Fusion community²⁰. The Fusion community wish to demonstrate making use of EOSC computational and storage resources for running containerised modelling applications (primarily HPC and HTC). This requirement derives from the fact that in digital infrastructures for research in fusion, local resources are not scaled for peak demand and the community wishes to use infrastructure available through the EOSC (and public cloud providers) as a scalable, non-vendor specific resource.

At a high level, this is an opportunistic use case where the community wishes to make use of any spare resources, and thus going through an ordering process would be non-optimal since the user would not know local resources are exhausted until they have submitted their computation tasks. It may be that some sort of framework agreement would be needed between the community and the providers to allow this opportunistic use. Different parts of the workflows may involve different computational requirements, from simple single core machines to many core/multi-node or GPU.

Storage and Data Management

Findability of language data collections with CLARIN, a European Research Infrastructure set up to provide access to language resources and tools for researchers who work with language data in the form of text, speech and mixed modalities. To be able to use language data, researchers need the support of a service platform that can turn linguistic content into a directly usable data source, via language-specific tools such as dictionaries or grammars, as well as generic services. CLARIN makes use of EOSC-hub solutions for compute and storage services. For example, the Virtual Language Observatory is a search tool that can be used to find relevant data by searching through metadata. Searching can be done on various facets, e.g. date, location, genre, collection. The primary users are researchers working in a specific domain for which the set of facets is optimised. Different instances for specific domains can be provided. Community data managers can use the tool to make their data available to a larger audience. B2FIND from EUDAT CDI is an example of a general metadata catalogue aimed at non-specific researchers for the discovery of CLARIN collection-level metadata.

¹⁹ <https://eosc-hub.eu/eosc-in-practice-wenmr>

²⁰ Source and credits: The CCFE: Fusion research, EOSC Portal Booklet, Nov 2018 (https://www.eosc-hub.eu/sites/default/files/EOSC_Portal_Booklet.pdf)

The precise makeup of the Shared resources is still under discussion, but EOSC-hub sees them as fitting into several broad categories. For example, access to large scale infrastructures for storage or processing of data could support a wide variety of communities, including the long tail of science, communities of practice, and short-term research collaborations who may have no easy access to internally managed infrastructures, as illustrated in the use cases above. This is complementary to the common services of the EOSC Service Portfolio, which are typically still more focused on a specific capability in the data lifecycle management process, rather than being the full general-purpose infrastructures imagined here.

Shared resources can potentially include access to datasets and other scientific products (e.g. software, applications and pipelines) otherwise not easily accessible in a sustainable coordinated manner through existing research infrastructures. Examples of such datasets include the Copernicus long term archive, Earth Observation data and genomic data archives, which cannot be accessed and processed in a scalable way at the researcher's in-house facilities. The Federating Core should provide access to such open science resources through the same portals and systems used to access other services. However, the datasets themselves may be mirrored, stored and managed in collaboration with the respective data providers, and could provide mechanisms of ingestion back to EOSC to allow users to share tools, software and data that are the output of their research. The added value for researchers here would be that they are exposed to such public good, open science datasets and can integrate them into their research more easily.

EOSC-hub also envisages that the Shared resources could include repositories for storage and processing of Digital Research Objects in support of open science. Due to the digitisation of science, it is now possible to collect together the full 'source' material for a piece of research: the source data, analysis pipelines, result data and scientific publication derived from it. These can be collected and stored together in a way that would support the traceability and reproducibility of science, as well as opening up new opportunities for promoting and supporting new research avenues. Infrastructures storing, managing and processing such objects could bring considerable benefits to the European Research Area.

The definition of Shared Resources needs to be evolved through a thorough and more complete analysis of use cases emerging from EOSC implementation projects and supporting initiatives.

2.3 EOSC in the Research Cycle

An illustration of the value the EOSC Federating Core and Service Portfolio could potentially provide is given in figure 2. Together, the EOSC Federating Core and EOSC Service portfolio will enable the full research data lifecycle management.

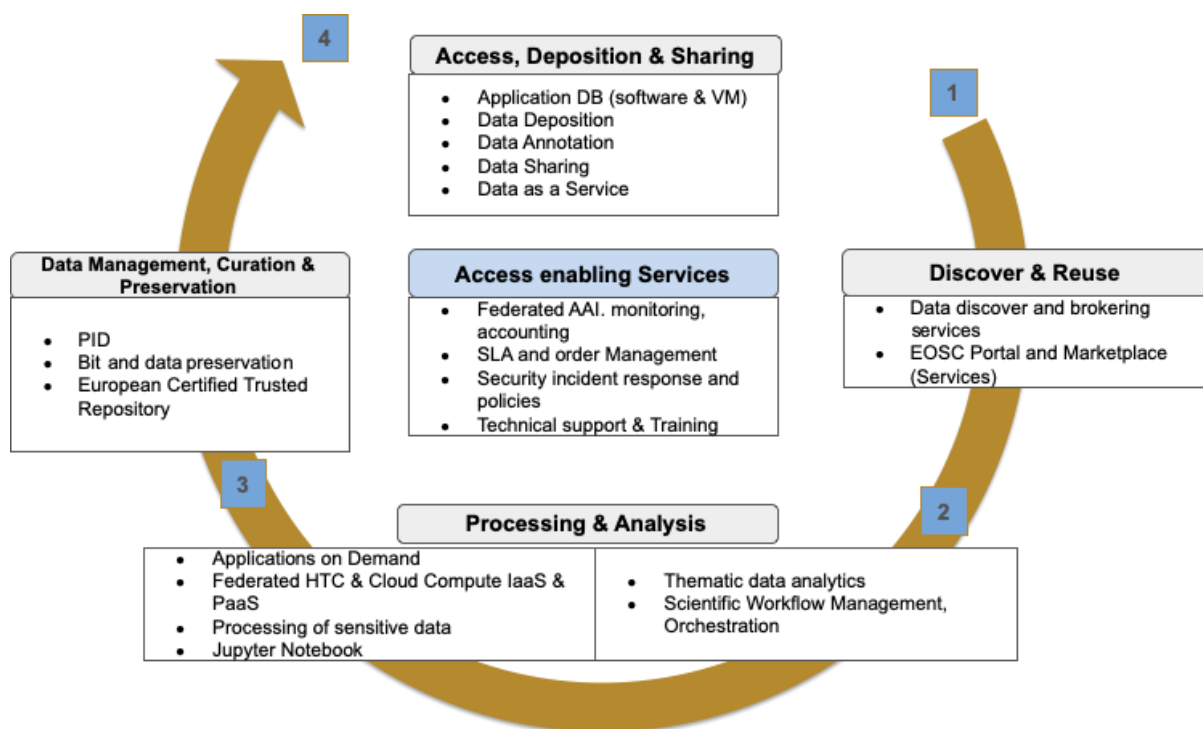


Fig 2: The role of the EOSC Federating Core services (in the middle) and various common and thematic services of the customer facing EOSC Service Portfolio in different phases of the research cycle.

3. Federating Core Services

An initial list and description of the services which may contribute the Hub Portfolio and Compliance Framework is provided below. The Federating Core structure and capabilities listed below are based on the operational experience of EOSC-hub project partners managing federating infrastructures, and the user requirements being collected from thematic service providers, EOSC-hub Competence Centres and the use cases emerging from the EOSC Early Adopter Programme involving research communities. We plan to review, complete and further expand the Federating Core description in collaboration with the EOSC community (e.g. EOSC stakeholders, EOSC governance and the implementation projects) as described in Chapter 5.

3.1 Hub Portfolio

Service Name, Description and Use Case
<p>EOSC Portal. The EOSC Portal provides a European-level delivery channel connecting the demand-side (the EOSC Customers) and the supply-side (the EOSC Providers) to allow researchers to conduct their work in a collaborative, open and cost-efficient way for the benefit of society and the public at large. In particular it delivers the following functions:</p> <ul style="list-style-type: none"> • Enable different kinds of users, with different skills and interests, to discover, access, use and reuse a broad spectrum of EOSC Resources (services, datasets, software, support, training, consultancy, etc) for advanced data-driven research • Support interdisciplinary research and facilitate Resource discovery and access at the institutional and inter-institutional level • Allow researchers and institutions to focus on value creation through sharing and reuse as opposed to duplicating Resources and increase excellence of research and European competitiveness • Improve the provisioning of access to integrated and composable products and services from the EOSC Catalogue • Facilitate the composition of services and products to support multi-disciplinary science for example with high-level community-specific interfaces for running workflows involving EOSC services • Help Providers gain additional insight into potential Users outside their traditional constituencies • Give Providers the possibility to offer Resources under homogeneous terms of use, acceptable use policies, and in different configuration options, so that Users are guided in the choice. <p>Use case. The Portal is particularly relevant to support on-demand access to EOSC through Business-to-User (B2U) and Business-to-Business (B2B) transactions.</p> <ul style="list-style-type: none"> • B2U is applicable for consumer-oriented Resources appealing to a large potential User pool. B2U transaction will address the digital needs of individual researchers and short- and medium-term research projects. Because of the large user base, B2U transactions will be possible for those Resources supporting automated or semi-automated provisioning, a short acquisition process, requiring a low-level of specialisation, and which can be easily compared and chosen without requiring expert support. • On the other hand, B2B applies for the acquisition of bespoke solutions and/or of large quantities of EOSC Resources involving potentially multiple Providers. B2B suits the needs

of research performing organisations and research infrastructures which need to cater for the long-term needs of a large pool of end users.

The EOSC Portal Concept 2.0²¹ provides extensive information on potential use cases and a participatory model for resource providers, which are provided with the choice of selecting different EOSC participation levels (e.g. Entry, Standard and High).

EOSC Helpdesk. The helpdesk is the tool that supports Incident and Service Request Management to restore normal/agreed service operation within the agreed time after the occurrence of an incident, and to respond to user service requests. The service works as a unified ticketing system, by connecting individual providers' helpdesks to the central helpdesk instance, offering a standalone service interface.

Use case. The helpdesk tool is necessary to support Incident and Service Request Management of the resources provided by EOSC. The helpdesk can be implemented as a distributed platform linking together the helpdesks of the suppliers offering resources to EOSC. The linking of existing helpdesks allows streamlining of support processes involving multiple suppliers, and in particular facilitates the work of the support teams that, through linking, are able to use existing in-house helpdesk tools.

EOSC Support Services. Users receive technical support, training and advice on specific policy and technical aspects.

Use case. One of the main EOSC pilot recommendations suggests that users receive support for cross infrastructure workflows, as these "frequently need to run across resources requiring integrated AAI, orchestrations tools and services, workflow languages and service interoperability." Support services are required to provide "easy access to large scale resources and services".

EOSC AAI. The EOSC AAI enables seamless access to research data and services in EOSC in a secure and user-friendly way.

The EOSC AAI follows the architectural and policy recommendations defined in the AARC project [[AARC-Community](#)]. As such, it enables interoperability across different SP-IdP-Proxy services, each of which acts as a bridge between the community-managed proxies (termed Community AAIs) managing the researchers' identity and the generic services offered by Research Infrastructures and e-Infrastructures (termed R/e-Infrastructures or Infrastructures). This is the "community-first" approach to the AARC Blueprint Architecture [[AARC-G045](#)], which enables researchers to sign in with their community identity via their Community AAI. Community-specific services are connected to a single Community AAI, while Infrastructure Services are connected to a single Infrastructure Proxy. Lastly, generic services may be connected to more than one Community AAI. Each Community AAI in turn serves as a bridge between external identity providers and the proxies to the e-infrastructure services. Specifically, Community AAIs connect to eduGAIN as service providers but act as identity providers from the services point of view, thereby allowing users to use their credentials from their home organisations. Complementary to this, users without an account on a federated institutional Identity Provider are still able to use social media or other external authentication providers for accessing services.

Research communities can leverage the EOSC AAI services for managing their users and their respective roles and other authorisation-related information. At the same time, the adoption of standards and open technologies, including SAML 2.0, OpenID Connect, OAuth 2.0 and X.509v3, facilitates interoperability and integration with the existing AAIs of other e-Infrastructures and

²¹ <https://wiki.eosc-hub.eu/display/EOSC/EOSC+Portal>

research communities.

Use Cases. Access to all EOSC shared resources and access enabling services (e.g. the Portal, the Helpdesk, EOSC data and compute and storage resource tier) will require federated authentication and authorisation. In addition, the Life Science Research Infrastructure cluster, as well as other research infrastructures from other scientific domains like Social Sciences and Humanities and Physics, have been piloting different solutions to get AAI as a managed service.

Monitoring. Monitoring provides the capability of checking the status of service end-point interfaces and aggregating such information for the production of service reports. In particular, it should provide a scalable framework for monitoring status, availability and reliability. It provides monitoring of services, visualisation of their status, dashboard interfacing, notification system and generation of availability and reliability reports. Third parties can gather monitoring data from the system through a complete API. A central deployment monitoring engine in EOSC can serve to reduce maintenance and integration costs.

Use case. Monitoring information supports Service Report Management, and is consumed to produce Service Reports, i.e. the documents that provide the details of the performance of a service against the service targets defined in service level agreements (SLAs) – often based on key performance indicators (KPIs). Typical users are the EOSC service suppliers.

EOSC Accounting. Accounting is about collecting, aggregating, storing and displaying EOSC resource usage data produced by the providers participating in EOSC, for example from the providers of Shared Resources. It gathers usage information from the individual resource providers and aggregates it centrally in a secure, GDPR-compliant manner. Accounting is necessary for providing control over resource consumption by the funders, and reduces the overhead of defining accounting information models, architecture and setup. Accounting is a key service of the EOSC federating core that will support its business models, and provides transparency on which resources are being used. The correlation of usage data to service identifiers, scientific product identifiers and user identifiers, supports the development of altmetrics that relate scientific impact to the extent a researcher and/or project has been embracing open science practices.

Use case. Accounting of resource usage is required for any EOSC customers (e.g. platform operator and research infrastructure managers) to get aggregated information on usage of scientific products and services used from the EOSC portfolio, to scale up the in-house infrastructure. Examples of such use cases are the Photon and Neutron research infrastructures, who plan to make use of EOSC shared resources to make data available for third-party exploitation when the data embargo period comes to an end, or the operators of exploitation platforms for Copernicus datasets as a service, aiming at accessing data for evidence-based policy making (e.g. in water management, forestry and agriculture).

EOSC Configuration Management Database (CMDB). The configuration database is an ITIL database used by an organisation to store information about hardware and software assets (commonly referred to as Configuration Items). This database acts as a data warehouse for the organisation and also stores information regarding the relationship between its assets. The CMDB provides a means of understanding the organisation's critical assets and their relationships.

Use case. The availability of an EOSC CMDB is relevant to EOSC shared resource suppliers, and is requested by the IT configuration management process. It allows the management of the provision of services owned and managed by the EOSC governance. It is envisaged that the management of resources published in EOSC just for the purpose of improving their discoverability, will be delegated to the respective providers and will not be registered in an EOSC

CMDB.
<p>Collaboration Software. Issue management and documentation co-development and sharing. Use case. Collaborative tools are typically relevant to the suppliers of EOSC shared resources for monitoring the progress of activities and actions, and for the sharing of information and documentation requested by IT service management processes.</p>
<p>Operations Portal. The Operations Portal refers to the set of control dashboards that support the work of EOSC infrastructure managers in charge of supervising the overall status, allocation and accessibility of the EOSC shared resources. It provides central operations management of federated resources. The Operations Portal offers a portfolio of management tools to support communications, customer relationship management, infrastructure oversight, and metrics gathering. Use case. The Operations Portal can support multiple service management activities like incident management and order management if used as a back-office tool of the EOSC Portal.</p>

3.2 Compliance Framework

Name, Description and Use Case
<p>Rules of Participation. The policies for users and suppliers to participate in EOSC. i.e. to become an active consumer of an EOSC resource, and to participate in the provisioning and sharing of services and scientific products through EOSC. They define the EOSC regulatory framework, which broadly speaking defines technical interoperability guidelines, FAIR guidelines, quality of service, and acceptable use policies to be observed by users and suppliers. Use case. The adoption of a single Acceptable Use Policy (AUP) for users is important to remove barriers in access, for example by requiring the acceptance of a single AUP applicable to all resources accessible through EOSC. From a supplier point of view, different sets of rules can be defined depending on their chosen level of engagement with the EOSC. For example, for suppliers choosing the Entry level, a minimum set of policies can be required, including compliance to a common service description template, the ability to publish service metadata according to agreed guidelines, and adherence to EU and applicable national regulations. For the Standard level, the Rules of Participation can be extended to require conformance to a given set of FAIR guidelines and interoperability standards. For the High level more stringent rules can be defined, for example the adoption of a common set of service management processes, and the guarantee of a minimum quality of service. The Rules of Participation address some of the EOSCpilot recommendations including: (1) "Addressing FAIR data principles on a community by community basis" (WP4 EOSCpilot recommendation²²) (2) "Support for cross infrastructure workflows" by ensuring adherence to common guidelines that are validated at service onboarding time, and periodically re-checked. (WP4 EOSCpilot recommendation)</p>
<p>Service Portfolio Management Tool (SPMT). The Service Portfolio management tool allows lifecycle management of the resources provided by EOSC. In particular, it aims at facilitating service management in IT service provision, including federated scenarios. SPMT represents a complete list of the services managed by a service provider; some of these services are visible to</p>

²² <https://www.eosc-pilot.eu/content/d44-consolidated-science-demonstrator-evaluation-report>

the customers, while others are internal. The service management system is designed to be compatible with the FitSM standard²³.

Use case. The tool is used by the team involved in the activities of the Service Portfolio Management process, which defines and maintains a service portfolio. A service portfolio is the entity that provides information such as the service value proposition, target customer base, service description, relevant technical specifications, cost and price, risks to the service provider, service level packages offered, etc.

EOSC Service Management System (SMS). Service Management System refers to the set of policies, roles and processes that control and support management of services within an organisation or federation, making sure the delivery is organised in a way that is controlled and repeatable. The Service Management System applicability is defined by its scope.

Use case. An EOSC Service management System is required for the provisioning of resources managed by EOSC. These resources can be supplied in a federated manner by various suppliers active at national and European level.

The provisioning of a Service Management System includes a number of processes which have been recommended by EOSCpilot leveraging the experience of various scientific demonstrators, such as:

- (1) Customer Relationship Management, to ensure a "close and continuous collaboration between EOSC and the active research communities" (WP4 EOSCpilot recommendation²⁴)
- (2) Incident and Service Request Management, to ensure the availability of "IT and domain aware experts to support researchers" (WP4 EOSCpilot recommendation)
- (3) "Make service descriptions semantically rich and easily discoverable: Integrate Service and Infrastructure Providers to deliver usable environments, and Provide integrated managed services" (WP4 EOSCpilot recommendation).

²³ <https://fitsm.itemo.org/>

²⁴ <https://www.eoscpilot.eu/content/d44-consolidated-science-demonstrator-evaluation-report>

4. Sustainability, Governance and Costing

4.1 Levels of EOSC Value-Add

It is expected that EOSC service providers will engage with the EOSC to differing degrees depending on the extent to which they wish or need to integrate their services with those of the EOSC – that is, the EOSC can act to varying degrees as a federator of existing systems. The offer and use of services through the EOSC are still at an experimental stage and definition of the EOSC federating core - in particular definition of its Shared Resources component – is a recent and ongoing activity, so of necessity models to support integration of services with EOSC are also still under definition. Three possible levels of EOSC value-add are described here by way of illustration. The levels of value-add are broadly based on and derived from the added-value roles described in EOSCpilot deliverable D5.1 “Initial EOSC Service Architecture”²⁵. They are:

Entry/Invisible Coordinator: resources from multiple suppliers would be available through the EOSC, but a resource would be natively delivered to EOSC users through its supplier’s specific delivery channel. The role of the EOSC would be limited to service advertisement and discovery.

Standard/Matchmaker: the EOSC would match a user request for a specific resource or functionality with the multiple EOSC service suppliers which can offer it. The request would be allocated to the “best option” based on defined criteria. The EOSC would provide service discovery but also the possibility of integrating with additional federated core services such as AAI, accounting, monitoring, helpdesk as a managed service and security coordination as a managed service. Service providers opting for this level of value-add would adapt part of their service management framework to EOSC.

High/One-stop Shop: the EOSC would play the role of a “resource integrator” for multiple EOSC service providers, to offer a complex bespoke solution to the user. The heterogeneity which may lie behind the solution would be hidden to the user. The EOSC would provide not only discovery and integration with federated core services, but also interoperability based on federation guidelines to allow integration of services. The EOSC would provide billing, monitoring and SLAs.

From the point of view of service provision, the three examples of value-add levels provided above are expected to require different mandatory functionalities from the EOSC Federating Core.

4.2 EOSC Access

In the context of the Research Infrastructure Charter for Access²⁶, three main modes of access and funding are envisaged. The choice of the applicable access models to EOSC is strategic and is influenced by its business model.

- Excellence-Driven Access: exclusively dependent on the scientific excellence, originality, quality and technical and ethical feasibility of an application, evaluated through peer review conducted by internal or external experts. This enables users to gain access to the best facilities, resources and services wherever located. In particular this access mode is suited for

²⁵ <https://eoscipilot.eu/sites/default/files/eoscipilot-d5.1.pdf>

²⁶ https://ec.europa.eu/research/infrastructures/pdf/2016_charterforaccessto-ris.pdf

the allocation of technically or economically scarce resources. This mode enables collaborative research and technological development efforts across geographical and disciplinary boundaries.

- **Market-Driven Access:** is defined through a contract between the User and the e-Infrastructure that will lead to a fee for the Access and that may remain confidential. This access mode is particularly suited for resources which can scale with usage demand. It should be noted that this type of access can apply to services offered by commercial providers and publicly funded ones.
- **Wide Access mode:** guarantees the broadest possible gateway to scientific data and digital services provided by the e-Infrastructure to users, wherever they are based. Typically, this mode is suited to resources which are not prone to congestion (i.e. which scale well with relatively flat funding levels). Adopting this mode maximises availability and visibility of the data and services provided.

4.3 Sustainability and Governance

In line with the HLEG recommendations, we propose the entire EOSC Federating Core to be entrusted to the governance of the EOSC. Funds should be made available to sustain the costs of federating existing digital infrastructures – leveraging existing investments, and expanding their capacity. Funding should come from member states, and multistate entities (like the EC, international research organisations etc.) to support Excellence-Driven and Wide Access models. Market-driven access will be sustained by paying customers, like research projects and commercial entities.

We propose the Hub Portfolio and Compliance Framework elements of the Federating Core to be part of the “Minimum Viable Ecosystem” to be accessible through the Wide Access mode to all suppliers participating in the EOSC federation. Therefore, we propose that the components or solutions which provide the Hub Portfolio and Compliance Framework services are governed by the EOSC. Their sustainability would therefore be synonymous with that of the EOSC itself.

We propose the Hub Portfolio and Compliance Framework elements of the Federating Core to be part of the “Minimum Viable Ecosystem” entrusted to EOSC and accessible through the Wide Access mode to all suppliers participating in the EOSC federation. These two components, central to the EOSC, are necessary to build the integration and interoperability framework across the services and scientific products that are currently provided through multiple different policies and access channels.

We expect that all three access modes - wide access, excellence-driven and market-driven – are likely to be supported by Shared Resources, depending on the user segment served and economic characteristics of the resource. Long-term funding to the Shared Resources will maximise the exploitation of shared scientific products and create economies of scale across the member states and international research communities. In particular, we propose:

- Wide access and excellence-driven depending on the amount of resources needed: for the long tail of science and application operators making use of the EOSC resource tier for research exploitation. This could entail a compensation mechanism to cover at least marginal and potential alternative costs where member state investments are used.
- Market-driven: for long-term research collaborations, projects and commercial exploitation. This requires research projects to have sufficient budget (i.e. purchasing power) and service providers to have mature accounting and invoicing capabilities.

Finally, we propose that the services of the EOSC portfolio are independently owned by their respective providers. These services are made discoverable and accessible through the EOSC portal, using EOSC as an “invisible coordinator”.

The proposed EOSC federating core funding model is in line with the HLEG report which recommends that the EOSC business models:

- “allow International, national and private funding entities to maintain key electronic data and software resources for the benefit of the scientific community”
- “enable reuse in situ of high value scientific digital objects, by which we mean data, software, metadata, workflows and other digital artefacts of scientific research.”
- “ensure that the private sector re-invests in R&D to stimulate innovation and create new markets”
- “create an EOSC MVE, rather than simply another set of digital silos, recipients of support would need to meet a set of technical and operational standards that would ensure that these resources (storage, computing, and higher order services such as software, pre-defined workflows, etc.) are accessible to scientists outside of the host institution and across Member States”.

4.4 Indicative Costing

The following table provides a **partial ballpark indication** of the cost of activities and services of the Hub portfolio and the Compliance Framework, looking at the annual allocation of efforts that are in the scope of the EOSC-hub. These costs provide an initial indication, **but they do not include costs of federating core services outside the scope of EOSC-hub or any costs in the resource tier**, and will vary substantially with any possible EOSC implementation. It should be noted that it is our plan to evolve the federating core description and update cost estimations in collaboration with OpenAIRE-Advance, FREYA, FAIRsFAIR, the RI clusters and national projects.

The activities included in such estimation are:

- Software maintenance and operations of services in the Hub portfolio, including 2 Million Euro currently allocated to the consolidation of the EOSC Portal.
- Advice, training and outreach (DMP training and training infrastructure, generic technical support, Digital Innovation Hub). While the costs of coordinating support services and of provisioning of a generic training infrastructure are approximately constant, the technical and policy areas covered and the provisioning of bespoke integrated solutions to specific user

groups greatly affect the costs of this service category. The figures provided in the table refer to the:

- provisioning of support and training through 8 Competence Centres,
- the operations of a Digital Innovation Hub for SMEs and industries that require a small amount of pledged capacity and no service customisation, and
- the provisioning of training and support on generic compute and data management technical solutions, and a selected group of nine thematic services.
- Maintenance and adoption of a Regulatory framework including policies and a subset of the service management system (SMS) processes needed to operate a federation of about 250 service providers. Both the regulatory framework and the SMS have a cost of maintenance and provisioning which only marginally increases with the number of providers participating in the federation.

Table 1. Annual costs of provisioning some of the federating core components according to EOSC-hub activity plan, the INFRAEOSC-06 budget, and the operational experience of the EGI Federation and the EUDAT CDI. The average PM rate used to estimate the annual cost in Euro amounts to 8,560 Euro/month.

Federating Core Component	Activity / service	Cost (Person Months / year)
Federating Tier (Hub portfolio) / Federating services (software maintenance)		
	Federating tool software maintenance (AAI, EOSC Portal and Marketplace capabilities, Integrated system and operations tools, Monitoring and Accounting, Helpdesk)	170
Federating Tier (Hub portfolio) / Federating tool operations (daily operations)		
	Daily running of the Hub portfolio technical infrastructure	218
Federating Tier (Hub portfolio) / Enabling services (Consultancy, Training, Support)		
	User and service provider engagement, training of generic and thematic services (9 instances), bespoke support to 8 user groups through Competence Centres, the Digital Innovation Hub and generic technical support to users	356
Regulatory Tier		
	Technical interoperability guidelines, rules of participations, data policies and the running of the Service Management System (Coordination, portfolio management, orer and customer management, IT security management)	148
TOTAL COSTS / year		
	Person Months	891
	FTEs	74,25
	Euro/year	7.624.287

It is important to note that costs in Table 1 are based on the current operational experience of running the EGI Federation and the EUDAT CDI. These provide an indication of the operational costs of the EOSC Federating Core, **excluding the shared resources component**, which will be determined by the amount of capacity and the set of scientific products to be discovered, accessed, and exploited in

EOSC. The cost analysis of the Federating Core should be updated once the list of capabilities to be provided has been integrated and updated with additional input from the EOSC community.

The EOSC governance is entrusted to select and continually evolve the set of federating core services, and to define the technical criteria and selection process that will allow the appointment of the most suitable suppliers. Actual costs in practice will be determined by the technical specifications of the federating core and the selected suppliers.

4.5 Access and Compensation

In current practice, the shared resource tier is typically not funded at the EU level but dependent on the allocation of member state investments to pan-European or cross-disciplinary collaborations. This practice relies on orchestration of resources and political mandates and a complex and lengthy resource allocation process. In the last two EC Research and Innovation Framework Programmes, the EC has experimented with different mechanisms to compensate service providers and overcome the above-mentioned dependencies, to allow more flexible and scalable resource allocation for pan-European or cross-disciplinary collaborations.

In particular the EC has promoted the Virtual Access (VA) scheme which aims to widen the use of existing services to new user communities and as such assumes the Wide Access mode. This mode limits the potential to exclude usage, effectively providing access on a first-come, first-served basis and rendering the mechanism unsuitable for scarce resources. VA may be well suited to an existing scientific database where use by one party does not exclude use by another, but not for providing storage or computing capacity.

VA allows service providers to recover the additional costs for on-boarding new user communities, such as training, but excludes any capital expenditure. Investments required to increase capacity to supply the additional demand thus cannot be recovered. The available capacity within the original (e.g. national, institutional or disciplinary) remit is therefore reduced or requires institutional funding to fund the capacity increase, which may not be politically viable. This limits the potential for capacity-bound services to scale with international demand, which may ultimately lead to a demand surplus.

Finally, VA requires service providers to account for the increase of usage and ensure all costs are identifiable and justifiable as direct costs. This means any use of shared infrastructures introduces high administrative burdens and associated costs, e.g. timesheets for administrative staff working on underlying infrastructures, creating an incentive to establish dedicated infrastructures to allow for easier accounting, causing fragmentation and significantly limiting the potential to seek efficiencies through economies of scale.

In summary the VA scheme neither allows capacity-bound services to limit usage nor to scale with demand, while bringing high and costly administrative burdens to provide the service to new user communities and producing incentives for fragmentation.

Any sustainable access and compensation scheme for the EOSC should take into account the economic characteristics of the resource being provided. The access mode applied should be appropriate to the size of the demand in relation to the scarcity of the resource. As a minimum, the mechanism should be able to compensate for the marginal costs of added usage with light administrative procedures, to

make it economically and politically viable for service providers to supply within the context of the EOSC.

A mechanism based on unit costs would allow the leveraging of accounting and transaction practices used for the Market-Driven access mode, creating a synergy between the different access modes at the funding level as well as achieving consistency in the compensation mechanisms of public and commercial sector service providers regardless of access mechanisms – the only difference being the funding source.

5. Concluding Remarks

The proposals in this document are provided by way of initial input to assist with the ongoing implementation of the EOSC. They will also be used towards defining a governance and sustainability roadmap within EOSC-hub, which will be ratified by the EGI, EUDAT CDI and INDIGO-DataCloud governance. Work will continue within the EOSC-hub project to define and implement the EOSC Hub, in collaboration with other EOSC implementation projects, and including OpenAIRE-Advance to deliver the JAM3.4 aligned roadmap(s) for service positioning and sustainability.

We welcome feedback. Input on the proposals contained in this Briefing Document will be collected from consultation with the EOSC stakeholder community. We propose this briefing paper to initiate a community activity aiming at discussing and evolving the Federating Core concept, This should be ultimately reflected in a community position paper which we envisage to include a mapping relating Federating Core capabilities to activities and services from the EOSC community (projects and supporting initiatives).