Solutions for a Sustainable EOSC

A FAIR Lady (olim Iron Lady) report from the EOSC Sustainability Working Group

Independent Expert Report

Solutions for a Sustainable EOSC

A FAIR Lady (olim Iron Lady) report from the EOSC Sustainability Working Group

Edited by: the EOSC Executive Board

November 2020

Table of Contents

ABSTRACT

This document explores possible means for sustaining the European Open Science Cloud beyond its initial phase which terminates at the end of 2020. This independent document builds on earlier strawman and tinman versions and the feedback received from the EOSC Executive Board, Governance Board as well as the European Commission (EC) and the stakeholder community on each version. It also takes into account the progress towards the EOSC goals as well as the outputs of commissioned studies. It considers the financing model, legal vehicle, governance structure under the planned European Partnership with the EC as well as the regulatory and policy environment of the EOSC. It recommends beginning with a first iteration to establish a Minimum Viable EOSC (MVE) addressing the needs of publicly funded researchers exploiting openly available data. Subsequent iterations expand the EOSC to address usage beyond openly available FAIR data and engage a wider user base including the public sector and the private sector.

Editors: Bob Jones, Maria Häll, Henriikka Mustajoki

MEMBERS OF THE SUSTAINABILITY WORKING GROUP

Rupert Lück, Executive Board, EMBL (Co-Chair)
Lidia Borrell-Damián, Executive Board, Science Europe (Co-Chair)

Bob Jones, EOSC Secretariat, CERN (Rapporteur and Main Editor)

Emanuele Barbarossa, DG RTD, European Commission
Jean-Yves Berthou, INRIA, France
Paolo Budroni, TU Wien Bibliothek, Austria
Ernest Cachia, University of Malta, Malta
Mirjam van Daalen, PSI, Switzerland
Claire Devereux, UKRI, UK
Franciska de Jong, CLARIN ERIC, Netherlands
Maria Häll, Swedish Research Council, Sweden
John Howard, University College Dublin, Ireland
Kristina Lillemets, HITSA, Estonia (replaces Maarja Adojaan)
Tamas Maray, KIFÜ, Hungary (replaces Lajos Bálint)
Norbert Meyer, Poznan Supercomputing & Networking Center, Poland
Patricia Mergen, Royal Museum for Central Africa, Meise Botanic Garden, Belgium
Panayotis Moschopoulos, DG RTD, European Commission
Liina Munari, DG CNECT, European Commission
Henriikka Mustajoki, Federation of Finnish Learned Societies, Finland
Alberto Pérez, Red.es, Spain
Federico Ruggieri, GARR, Italy
Stelios Sartzetakis, Athena Research and Innovation Center, Greece (replaces Afrodite Sevasti)
Roar Skålin, The Norwegian Meteorological Institute, Norway
Gintare Tautkeviciene, Kaunas University of Technology, Lithuania
Klaus Tochtermann, ZBW - Leibniz Information Centre for Economics, Germany

Christos Marinos-Kouris, EOSC Secretariat (replaces Panoraia Spiliopoulou)

FOREWORD

The Sustainability Working Group held its first meeting in July 2019 and a Strawman version of this document was distributed in early September in order to gather initial feedback from the Governing Board, Executive Board and projects contributing to the EOSC. This feedback was analysed and contributed to the development of the Tinman version that was distributed in December 2019. By February 2020, more than 30 projects, organisations, board members, working groups and individuals provided written feedback on the Tinman document. We would like to thank all those that took the time to provide detailed feedback on the Strawman and Tinman versions.

This FAIR Lady (olim Iron Lady) version of the document takes into account the analysis of the Tinman feedback and on-going developments by the EOSC-related projects and other Working Groups as well as the results of several studies commissioned by the EOSCsecretariat.eu project:

- Legal and strategic advice on EOSC legal entity membership and governance structure: performed by Kellen Group

- EOSC core operational costs: performed by AcrossLimits and Boundaryless

- Strengthening of the EOSC Risk Governance through the implementation of an effective risk management system: performed by AON Hewitt

- Expanding EOSC: Engagement of the wider public sector and private sectors in EOSC: performed by Industry Commons Foundation

- EOSC Working Group Landscape, Assessment Country Sheets Analysis: performed by the Digital Curation Centre

- Fair Forever 2.0: Long Term Data Preservation Roles and Responsibilities – An Assessment Proposal for the EOSC Sustainability: performed by the Digital Preservation Coalition

These studies provided important insights for the planning of EOSC and we are grateful to the organisations and companies for their excellent work.

Rupert Lück, Executive Board, EMBL (Co-Chair),
Lidia Borrell-Damián, Science Europe (Co-Chair)

Acknowledgements

To the memory of our colleague and friend Dr. Lajos Bálint, who passed away suddenly in December 2019.

We would like to thank the following people for their contributions to this document and the efficient organisation of the Working Group:

Jessica Klemeier, EMBL
Marie Timmermann, Science Europe

INTRODUCTION

The European Open Science Cloud (EOSC) federates existing and emerging data infrastructures with the objective of offering a virtual environment in Europe to share and re-use research data across borders and disciplines. EOSC is expected to serve approximately 2 million researchers in Europe and progressively expand its user base to include the wider public sector and the private sector (business organisations).

An EOSC that offers added value to researchers was taken as a starting point with its scope as described in the Strategic Implementation Plan[1]:

> *"the EOSC should be a federation of existing and planned research data infrastructures, adding a soft overlay to connect them and making them operate as one seamless European research data infrastructure."*

Building on existing research data infrastructures, EOSC will grow through a series of iterations facilitating the interaction of existing and new infrastructures and supporting their move towards openness. Each iteration will enable greater functionality for a wider user base and satisfy a broader range of use-cases.

All relevant research and innovation actors, including scientific communities, research institutions, learned societies, community fora, national and European Infrastructures (generic or thematic), funders (public or private) and the private sector (including data and journal publishers) are seen as stakeholders of EOSC. The primary stakeholders are researchers as end-users, service providers (including data and compute services), as well as research funders and the EC. As defined in the EOSC glossary[2] end-users are researchers, including citizen scientists[3], that consume resources via services. Gradually, the EOSC user base will be expanded to the public and private sectors, creating solutions and technologies that will benefit all areas of the economy and society[4].

A key goal of EOSC is to shift the research system towards an Open Science model[5] as many European countries are implementing national programmes that are aligned with the European Commission Recommendation (EU) 2018/790 of 25 April 2018 on access to and preservation of scientific information[6].

Similarly, the EOSC could support the EU Member States in implementing the Open Data and Public Sector Information Directive on open data and the re-use of public sector information[7], which Member States have to transpose by 16 July 2021 and that includes publicly funded for research data[8].

---

1 European Open Science Cloud (EOSC) strategic implementation plan, 2019, ISBN 978-92-76-09175-2, doi:10.2777/202370 KI-03-19-507-EN-N

2 EOSC Glossary, September 2020, https://docs.google.com/document/d/1rA3XNw-ORrDzBUyZwOolacTOOBpjOaxWmFdg_EEZcpA/edit#heading=h.wzabkxo6b618

3 Citizen Science is "*scientific work undertaken by members of the general public, often in collaboration with or under the direction of professional scientists and scientific institutions*" https://ec.europa.eu/research/openscience/pdf/citizen_science_recomendations.pdf. When Citizen Scientists are acting as agents of an organisation or Institution, they fit the EOSC End-User definition; when not affiliated with an Institution, Citizen Scientists are defined as 'Consumers'.

4 European Cloud Initiative - Building a competitive data and knowledge economy in Europe (COM (2016) 178 final

5 https://ec.europa.eu/digital-single-market/en/open-science

6 https://www.eosc-portal.eu/sites/default/files/CELEX_32018H0790_EN_TXT.pdf

7 https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=uriserv:OJ.L_.2019.172.01.0056.01.ENG

8 In the sense of the directive, 'research data' means documents in a digital form, other than scientific publications, which are collected or produced in the course of scientific research activities and are used as evidence in the research process, or are commonly accepted in the

There needs to be political will not just towards the notion of open research, but also to fund the means to perform open research in practice. Governments and research funders have to be in a position to supply such funding, in accordance with a jointly agreed cost sharing model, and to coordinate their approach, policies and legislation.

It is essential that existing data infrastructures as well as the underlying high-capacity network that interconnects research, education and innovation communities, continues to be funded. As highlighted in the report entitled Guidelines on Cost Estimation of Research Infrastructures[9], data management including the tasks of making data FAIR and ensuring EOSC compliance can result in significant costs that are frequently not included in research budgets.

For EOSC to be a success, it must be widely adopted by researchers across national and discipline boundaries. This implies EOSC must provide access to services that allow researchers to pursue their research activities more effectively through faster and seamless sharing of publications, data, software and other digital research outputs. The study on EOSC core operational costs highlights that researchers seek, for example, a better reputation, a place to publish their outputs and be recognised and rewarded, finding opportunities for a career, or to become expert in a specific field and guide others with similar needs and issues. Training and co-research opportunities are also perceived as strong value gains. EOSC should capture this potential and provide services to support these needs of researchers.

The approach taken in this document is to place the researcher as an end-user at the centre and determine what added-value the EOSC will bring to them. While the services to be provided *via* EOSC to researchers are expected to be *free at the point of use[10]*, they are not without significant cost to build, promote, maintain and operate. Consequently, EOSC must also provide added value to stakeholders other than researchers, notably data and service providers as well as funding organisations, by delivering more and better science through open and collaborative knowledge sharing.

Therefore, the basic condition of success in ensuring EOSC sustainability is performance: how EOSC as an ecosystem operates and how the resources it federates are used and acknowledged by researchers. The sustainability of EOSC depends not only on sound business models encompassing the financial, legal and governance aspects to create added-value for the stakeholders but also on the incentives and rewards for researchers as discussed above that encourage them to participate in a culture of sharing the results of their research. Without such incentives and rewards it is possible that the uptake of the EOSC could be stymied by lack of engagement from researchers.

Finally, it is very important for the EOSC initiative to be considered inclusive and supplement the activities of mature and well-structured research communities that provide thematic services rather than competing with them, while at the same time respecting national strategies, business models, and relations between different stakeholder groups.

---

research community as necessary to validate research findings and results. Research data includes statistics, results of experiments, measurements, observations resulting from fieldwork, survey results, interview recordings and images. It also includes meta-data, specifications and other digital objects.

9 Guidelines on Cost Estimation of Research Infrastructures, StR-ESFRI Study, August 2019, https://www.esfri.eu/sites/default/files/StR-ESFRI2_STUDY_RIs_COST_ESTIMATION.pdf

10 *Free at the point of use* does not imply *Free of charge*. *Free at the point of use* means the end-user does not pay directly for the service when it is delivered but their consumption will be paid for by other means. For example, an end-user would not need to use a credit card to pay for a service but their employer may receive an annual bill from the service provider, or the employer has arranged a suitable subscription.

EOSC is a complex undertaking that engages multiple stakeholder groups with different perspectives and concerns. The EOSC vision is not limited to linking datasets, federating infrastructures, or aligning policies; it starts by linking multiple stakeholders throughout the data lifecycle and across the European research ecosystem. The feedback received to earlier versions of this document highlight a divergence of objectives, priorities and interests across and within stakeholder groups concerning the benefits and costs of EOSC as well as the disruption it may cause to existing structures, decision making processes and funding models. The planning of EOSC must acknowledge that such tensions exist and work to reconcile them if EOSC is to achieve its goals. The iterative approach as outlined in this document is intended to progressively build trust and resolve conflicts between the stakeholders while acknowledging that not all the solutions and answers are known today. For this approach to be successful there must be an objective assessment of what has been achieved with each iteration and the experience gained taken into account in preparing the future. The recent creation of the EOSC Association is an important achievement and the Association should strive to provide a forum where the stakeholders can resolve their differences and collectively confirm their commitment to achieving the objectives of EOSC.

1 FIRST ITERATION - A MINIMUM VIABLE EOSC (MVE)

The objective of the first iteration is to establish a Minimum Viable EOSC (MVE) such that it will enable the federation of existing and planned research data infrastructures for the benefit of publicly funded researchers accessing openly available FAIR data. The MVE is necessary in order to establish an initial implementation of EOSC that brings value to users beyond their current use of infrastructures.

The main focus and value of the MVE is to connect disciplinary infrastructures and research data to enhance disciplinary, cross-disciplinary and transnational research, leading to new scientific discoveries and new insights for society. Hence, the federation of research data infrastructures *via* the thematic clustering of Research Infrastructures, e-Infrastructures and regional projects[11] is seen as a critical first step to promote early successes for EOSC, and as an enabler to help address society's current and future global challenges.

This Graph implies that federated data is the glue between EOSC Core and Exchange

The MVE includes EOSC-Core and EOSC-Exchange, described below, that work with the FAIR datasets to be federated via EOSC.
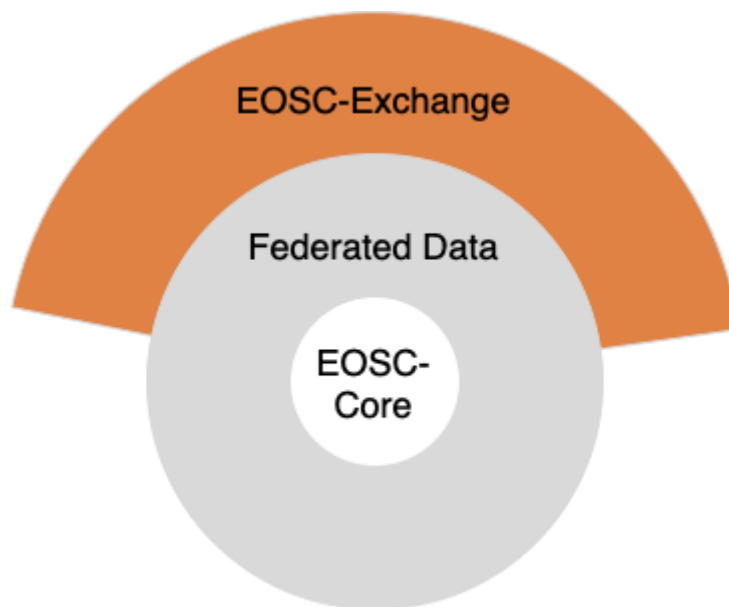
Figure 1: Schematic representation of key elements of the Minimum Viable EOSC

1.1 EOSC-Core

The *EOSC-Core* provides the functionality that is required to enable open science practices to occur across domains and countries according to the EOSC interoperability framework[12]. To facilitate long-term sustainability, the EOSC-Core is based on FAIR data principles and includes the minimum set of components necessary to provide the means to discover,

---

11 EOSC-related cluster and regional projects 5a: ENVRI-FAIR, EOSC-Life, ESCAPE, PaNOSC, SSHOC and 5b: EOSC-Pillar, EOSC-Nordic, EOSC-synergy, ExPaNDS, NI4OS-Europe

12 EOSC Interoperability Framework (v1.0), 3 May 2020, Draft for community consultation

share, access and re-use data and services. Its contents promote the adoption of an open research culture and must be maintained over the long term. Specifically:

- A standard mechanism for naming and locating data and services

- A mechanism for discovery of and access to data and services across the federated EOSC ecosystem.

- A common framework for managing user identity and access

While the *EOSC-Core* enables users and machines (preferably via open Application Programming Interfaces) to discover, share, access and re-use resources it is the services federated *via* EOSC-Core that actually transfer, store, process or preserve research data.

The initial implementation of *EOSC-Core* will be based on the widely used production quality components already deployed by the EOSC-related projects and communities to provide the following functionality for which more details are provided by the outputs of the Architecture[13] and FAIR[14] Working Groups:

- **A shared open science policy framework,** which effectively embeds a data compliance framework for open / FAIR data. It defines and applies the rules of how the data elements are published, shared and re-used.

- An instantiation of the **EOSC interoperability Framework** including:

  o **Authentication and Authorization Interoperability (AAI) framework,** a trust and identity service for researchers to seamlessly access any EOSC resource[15]. The AAI framework implements the AARC blueprint architecture[16] to provide a set of interoperable building blocks for international research collaborations. The Core includes those elements that provide identity inter-federation for establishing trusted communications between identity providers and service providers. Community-specific AAI services are necessary components of the EOSC AAI architecture, but they are not part of EOSC-Core.

  o **PID**: Services to generate, resolve and validate persistent identifiers (PID)

  o **An interoperable metadata framework,** for ensuring openness and interoperability across disciplines while respecting privacy and security (copyright status, disclosure limitations, patents pending, other IPR on the datasets or workflows, the existence of personal data, designation of data as PSI, etc.). Note that such a means of enabling interoperable metadata is a high priority for EOSC and is currently not addressed by the service providers consulted during the EOSC core operational costs study.

  o **Data access framework**, whose primary role is to offer data as a service. It enables open interfaces where data consumers (users and machines) are able to discover and use data.

  o **Service management and access framework,** whose role it is to provide a consistent and agreed upon understanding of e-science services: what they offer, which

---

13 https://www.eoscsecretariat.eu/working-groups/architecture-working-group

14 https://www.eoscsecretariat.eu/working-groups/fair-working-group

15 EOSC Resource extract from definition in the EOSC Glossary: EOSC Resources include services, datasets, software, support, training, consultancy or any other asset. https://www.eosc-portal.eu/glossary

16 https://aarc-project.eu/architecture/

science problem they address, what is their operational capacity, how they are accessed, who pays for them.

o **An open metrics framework,** which sets the rules (usage, performance, value for money, user satisfaction) for the assessment of EOSC elements, i.e., policies, access framework, services, data, business, funding and usage models. This should include elements to facilitate the incentives and awards mechanism for researchers, as recommended by the EC HLEG on Next Generation Metrics and the EOSC Pilot policy group[17].

• **Security policies and procedures** to ensure consistent and coordinated security operations across the federated services. This will include incident response policies and a service request and problem management scheme.

• **Operational support services** for EOSC-Core and made available to those federating services connecting to the EOSC-Core. Support services related to the individual services accessible via EOSC-Exchange or related to disciplinary data centres are not part of EOSC-Core.

• **Web-portal** with data and contents in multiple formats as well as supply and demand facing services providing for accessing the EOSC resources. It is expected that other web-portals will also exist and be developed outside of the EOSC-Core.

Building on the items listed above, EOSC-Core will provide the means to operate EOSC-Exchange as a digital marketplace of resources for publicly funded researchers. It will also include a collaboration and communication service (organisational) and a messaging service (technical) that facilitate the interoperability of Core services.

The MVE will be considered operational when key datasets and services provided by EOSC related projects[18], notably the thematic[19] and regional[20] cluster projects, will be exploitable by end-users who are external to these projects. The importance of verifying the MVE by using progressively more sophisticated use-cases proposed by research communities has been highlighted by a number of the stakeholders and studies mentioned in the forward section.

1.2   EOSC-Exchange

EOSC-Exchange is a digital marketplace that builds on the *EOSC-Core* to offer a progressively growing set of services exploiting FAIR data and encouraging its reuse by publicly funded researchers. It is expected that services, such as those that store, preserve or transfer research data as well as those that compute against it, will be made available via *EOSC-Exchange*.

Participation in *EOSC-Exchange* as a service provider is without registration fee but service providers will be required to conform to predefined Rules of Participation. While the technical requirements for participation in *EOSC-Exchange* will be the same for all services, there may be differences in the legal and policy requirements for freely available and payment-based services. The process for provisioning services from publicly funded operators will differ from the process of provisioning services from commercial providers.

---

17 Next-generation metrics: Responsible metrics and evaluation for open science, 2017, ISBN 978-92-79-66130-3 doi:10.2777/337729

18 https://www.eoscsecretariat.eu/eosc-projects-list

19 https://www.eoscsecretariat.eu/communities/EOSC-ESFRI

20 https://www.eoscsecretariat.eu/communities/eosc-regional-projects

1.3   Federated Data and Services

EOSC is foreseen as a federated structure and the Landscape Working Group of the EOSC Executive Board has surveyed and documented[21] the landscape of infrastructures, initiatives and policies across Europe related to the development of the EOSC. Information has so far been collated on 47 Member States and Associated Countries.

The EOSC-Core will provide a shared Open Science policy framework for open and FAIR data. Within this context an objective of the Landscape Analysis Report from the Landscape Working Group was to identify to what extent EU Members States (MS), Associated Countries (AC) and Other Countries (OC) are ready to join the EOSC. The analysis report recognized that there is currently no formal definition of EOSC readiness for countries. However, there is general agreement that a number of policies related to Open Science, Open Access, Open Data and Open Learning should be in place at the national level for countries to foster an ecosystem that will enable the EOSC vision. Based on this understanding and based on the material compiled based on the country sheets the readiness of MS, AC and OC was assessed in the Landscape Analysis Report. The main finding of the report is that all European countries show a fairly high degree of EOSC readiness, in the sense that they have (or are in the planning phase of) national policies for Open Science, Open Access, Open Data and Open Evaluation. The only exception is Open Learning, for which the majority of MS and AC have no policy.

For the further development of the landscape analysis several recommendations are listed in the report. For example, it is necessary to standardize the underlying database (in particular the country sheets) and to enrich it with structurally comparable information. Also, a selection of standard policy elements should be collaboratively agreed for future policy landscape analysis. Future Options result from the validation workshop of the landscape report. In this context, an indicator system is particularly important with the help of which the development of the readiness can be monitored. This requires the development of a common, coordinated set of indicators. It is of particular importance to create incentives that ensure a regular update of the indicators. Issues of concern are the security of the data to be collected and limited staffing.

1.3.1   Prototype EOSC Catalogue and Marketplace

The offer of services and resources is currently managed via the prototype EOSC Catalogue and Marketplace[22] operated by the EOSC-hub project. The prototype EOSC Portal[23] serves as an example entry point to EOSC services and resources from many domains by enabling users to access and request services supplied at institutional, national and regional levels enabling them to process and analyse data in a distributed environment[24]. In order to develop a rich platform offering a wide range of services and resources, EOSC requires the participation of service providers. Services and resources are provided and maintained by different providers under a variety of licenses and access requirements (such as: accessible by users outside its original community; described through a common template focused on value proposition and functional capabilities; at least one service instance is running in a production environment available to the user community; publication of research data is FAIR; release notes and sufficient documentation are available; helpdesk channels are available for support, bug reporting and requirements gathering.)[25] As documented[26] by

---

21 Landscape of EOSC related infrastructures and initiatives, Report from the EOSC Executive Board Working Group (WG) Landscape, 10.08.2020

22 https://marketplace.eosc-portal.eu/

23 https://www.eosc-portal.eu/

24 https://www.eosc-portal.eu/services-resources

25 https://eosc-portal.eu/for-providers

26 Landscape of EOSC-Related Infrastructures and Initiatives, August 2020, doi: 10.2777/342650

the Landscape working group, the catalogue currently gives access to resources offered by 73 Service/Resource Providers and Aggregators[27]:

- 254 services;

- 4.4M datasets;

- 141,000 software and applications;

- 34.6M publications;

- and 3M other research products;

The resources currently offered via the current EOSC Catalogue and Marketplace are only a fraction of those identified by the Landscape Working Group. This indicates an opportunity for EOSC to expand and the need to accelerate the on-boarding activities. The resources and thematic services identified as ready for integration by the Research Infrastructure clusters should be prioritised for on-boarding because they represent high-quality resources valued by research communities and the Research Infrastructures have demonstrated their eagerness to contribute to EOSC and have already established communication channels, and at cluster-level best practices are being developed to ensure the interoperability of federated services and data that is essential as highlighted by the Interoperability Task Force of the EOSC FAIR Working Group.[28]

1.3.2   European COVID-19 Research Data Platform

The European COVID-19 Research Data Platform[29] has been running for approximately 6 months with more than 78,000 unique users and 2.7 million requests served from over 170 countries. The continuous cooperation on the platform amongst the EU, it's Member States and the wider EOSC-related stakeholder communities has the potential to set the scene for the future development of EOSC and the resulting lessons learnt should be taken into consideration.

1.3.3   Research Infrastructures

Research Infrastructures are key stakeholders in the EOSC ecosystem, both as data producers and service providers. They foster the definition, implementation and further development of advanced solutions for the effective provisioning and use of high-quality scientific data, with effective metadata descriptors, ease of access, interoperability and reusability, implementing the FAIR principles. EOSC provides a means to increase data sharing and reuse beyond RIs, since the EOSC engages not only RIs but also data, users and service providers from different national and regional backgrounds. As highlighted in a recent paper on the state of FAIRness in ESFRI Projects[30], the FAIR principles are recognised as important, but taking all necessary steps to ensure resources become FAIR will require investment and convincing all actors to change their practices will take time. EOSC must support the research communities in the process of making their resources FAIR by providing tools, training and expertise. This support needs to be accompanied by alignment of policies across Member States that promote the adoption of FAIR and are linked to research funding streams.

---

27 https://catalogue.eosc-portal.eu/home

28 EOSC Interoperability Framework (v1.0), 3 May 2020, Draft for community consultation

29 https://www.covid19dataportal.org/

30 https://doi.org/10.1162/dint_a_00045

1.3.4 Access and license policies

A political goal of EOSC is to overcome existing national and discipline fragmentation in order to promote open research across Europe and consequently the *MVE* should be as widely used as possible. However, it is recognised that controlled (authenticated and possibly authorized[31]) access may be required in order to respect ethical legal, social or commercial aspects and that the access policy for a resource may change during the research lifecycle. Such an access policy choice is a decision to be made by the resource provider but usage metrics should be tracked so that an impact assessment of EOSC can be made.

Licensing policies can also affect the adoption of the MVE and its impact. The EOSC-Core should have clearly defined requirements on the licensing policies of its components and their interfaces to ensure they remain openly accessible and cannot be controlled by a dominant party. Licensing policies for the federated infrastructures and contents of EOSC-Exchange can be more tolerant in order to facilitate participation in EOSC but should still adhere to the FAIR principles.

1.3.5 Long-term access to data

Long-term open data archives and preservation services are required to enable a sustainable EOSC and the sustainable access to data. Data preservation not only refers to the long-term storage of data, but also includes ensuring the preservation and maintenance of data, as well as its context, understandability, interpretability, authenticity and integrity. Preservation of scientific data requires data curation by IT experts as well as the active involvement of documentalists and scientists to guarantee long term usability of research data. As a result, data preservation on large scales is often organised by long-term international projects and organisations making use of services provided at institutional, regional and national levels. Throughout the Member States (MS) and Associated Countries (AC), there is a large variety of data processing services - from local, regional, and national services to international services. The EOSC Executive Board recognises that the availability of long-term data preservation services represents an important added-value for EOSC but responsibility for the curation and management of datasets must remain with the communities. **The interim findings of the FAIR Forever study by DPC[32], noted that digital preservation is not explicit in the context of EOSC and the roles, responsibilities and accountability for digital preservation are currently not clearly defined.** The general sense is that researchers are responsible for data creation and management until they are stored or transferred to institutional systems or disciplinary or national infrastructures. The extent to which institutions have been given or taken explicit responsibility for preservation is unclear, assuming even that they have the capability to deliver. The concept of data stewardship is present which implies preservation, but is more often seen as an ambassadorial role, between the researcher and other institutional departments and staff such as the computing services, institutional repositories, libraries or archives. **Clearer roles and responsibilities are needed, including the assessment of capability as well as functions, salaries and funding streams for preservation.** Tied to the roles/responsibilities challenge is a lack of clarity on preservation responsibility. There is an urgent need for consideration of strategy that delineates responsibilities and leadership for preservation.

---

31 Authentication verifies you are who you say you are while Authorization decides if you have permission to access a resource. If the policy of a service provider is to allow open access then Authorization may not be required to access their services.

32 FAIR Forever Updated Interim Statement Report, October 2020, Amy Currie and William Kilbride

1.3.6   Connecting EOSC and EuroHPC

EuroHPC[33] is developing a pan-European supercomputing infrastructure. It will permit the EU and participating countries to coordinate their efforts and share resources with the objective of deploying in Europe a world-class supercomputing infrastructure and a competitive innovation ecosystem in supercomputing technologies, applications and skills. The EC has proposed a new Regulation for the European High-Performance Computing Joint Undertaking[34] to maintain and advance Europe's leading role in supercomputing and quantum computing and enabling an investment of €8 billion in the next generation of supercomputers. The EuroHPC Joint Undertaking is a legal and funding entity with the aim of developing a pan-European supercomputing infrastructure and supporting research and innovation activities by developing a European supercomputing ecosystem, stimulating a technology supply industry, and making supercomputing resources in many application areas available to a large number of public and private users.

Research data is growing at an unprecedented rate and science demands to address global challenges such as climate change are driving the convergence of HPC, big data and artificial intelligence (AI). Enabling the interaction between EuroHPC and the EOSC will provide a continuum of data and computer services to research communities to help them address such challenges.

The EC has allocated €77.5 million of funding to ten HPC European Centres of Excellence (CoEs) in areas such as engineering, environmental science, renewable energy, materials modelling and design, molecular and atomic modelling, big data and global system science, bio-molecular research, and are contributing tools to optimise HPC applications performance. Collaboration between the HPC CoEs and the EOSC communities is important. EOSCsecretariat.eu has investigated[35] the current landscape of European CoEs for HPC in relation to the EOSC. There are already examples of HPC software and technologies featured in the EOSC ecosystem, and CoEs are eager to ensure their services are accessible via the EOSC catalogue. Furthermore, in line with EOSC strategic priorities, the CoEs uphold FAIR data principles. The CoEs have also highlighted areas where collaboration can take place on specific topics in the future including finding and registering FAIR data across repositories. The relationship between EOSC and EuroHPC was raised several times during the EOSC governance symposium 2020[36] and it was suggested to organise a workshop on the subject in a similar manner to the 2nd ESFRI RIs - EOSC workshop[37] that addressed the connection of ESFRI Research Infrastructures (RIs) to the EOSC.

1.3.7   EOSC in a global setting

The EOSC will be European and open to the world, reaching out over time to relevant global research partners and initiatives so that by 2027 there can be alignment and interoperability of relevant policies and infrastructures necessary to promote Open Science globally. Coordination fora including UNESCO[38] and the G7 Open Science Working Group[39]

---

33 https://eurohpc-ju.europa.eu/

34 https://ec.europa.eu/commission/presscorner/detail/en/ip_20_1592

35 Advancing engagement between HPC Centres of Excellence & EOSC (https://doi.org/10.5281/zenodo.3727821)

36 https://www.eoscsecretariat.eu/eosc-symposium-2020

37 https://www.esfri.eu/esfri-events/2nd-esfri-ris-eosc-workshop-research-infrastructures-shaping-eosc

38 UNESCO Recommendation on Open Science https://en.unesco.org/science-sustainable-future/open-science/recommendation

39 Within the G7 ministerial framework, a specific Open Science configuration has been established by the G7 in 2016 in Tokyo to share expertise, exchange best practices and develop synergies on Open Science paradigms in G7 members and beyond.

at the policy level, as well as COAR[40], CODATA[41], RDA[42] and WDS[43] provide an environment where the different layers of interoperability (legal, organisational, semantic and technical[44]) can be discussed with partners from around the world. There is a clear willingness to collaborate and it is expected that the first agreements could be brokered by coordination fora during the first iteration of EOSC. The observer status of the EOSC Association will offer the opportunity to organisations from other regions to participate in EOSC activities. The RISCAPE analysis[45] has provided new insights on the global landscape and the operations, services and organisation of RIs globally. The analysis highlights possibilities for collaboration between European and international facilities across seven research domains as well as e-infrastructures and data infrastructures where international connections are already well established. Consequently, there are many opportunities for collaboration particularly in the context of alignment of solving global challenges, global service access, increased efficiency and scientific excellence. The recent collaboration[46] between researchers in Europe, Australia and USA to perform COVID-19 viral genome analysis gives a clear example of why reproducible research performed globally in an open and transparent manner is important. A policy for international collaboration, potentially based on reciprocity of access and aligned with Horizon Europe rules on international cooperation, needs to be developed by EOSC via the collaboration fora.

1.4 Funding scheme

As noted in the draft EOSC partnership proposal[47], it is not easy to assess accurately the future investments necessary to implement the EOSC and achieve its objectives. EOSCsecretariat.eu sponsored studies[48] and EOSC related projects[49] have studied business models for EOSC and come to similar findings as described below.

A unique added value of the EOSC is its ability to enable researchers to reuse data and services through the same portal and this can only be achieved by bringing together all the elements of the MVE. Consequently, looking for sustainability only in a subset of the ecosystem would be a high-risk strategy and a missed opportunity to pursue the value driven approach typical of platforms that has led to their fast growth in terms of impact.

1.4.1 Funding international research

The EOSC core operational costs study[50] and the use cases examined by EOSC-hub[51] highlighted the fragmented and complex nature of the European research funding landscape and the associated difficulties involved in attempting to provision services across

---

40 Confederation of Open Access repositories (https://www.coar-repositories.org/)

41 CODATA, the Committee on Data of the International Science Council (http://www.codata.org)

42 Research Data Alliance (http://www.rd-alliance.org)

43 World Data System (https://www.icsu-wds.org/)

44 Layers of the interoperability model defined in the European Interoperability Framework, European Union, 2017, doi:10.2799/78681

45 INTERNATIONAL RESEARCH INFRASTRUCTURE LANDSCAPE 2019, A European Perspective; DOI 10.5281/zenodo.3539254; https://riscape.eu/riscape-report/

46 No more business as usual: agile and effective responses to emerging pathogen threats require open data and open analytics; Galaxy and HyPhy developments teams, Anton Nekrutenko, Sergei L Kosakovsky Pond ; ioRxiv 2020.02.21.959973; doi: https://doi.org/10.1101/2020.02.21.959973

47 https://ec.europa.eu/info/sites/info/files/research_and_innovation/funding/documents/ec_rtd_he-partnership-open-science-cloud-eosc.pdf

48 'EOSC core operational costs' and 'Strengthening of the EOSC Risk Governance through the implementation of an effective risk management system' funded via the co-creation budget of the EOSCsecretariat project

49 EOSC-hub project

50 https://eoscsecretariat.eu/news-opinion/results-open-call-eosc-core-operational-costs

51 EOSC-hub Briefing Paper – Provision of Cross-Border Services, EOSC-hub project, version 1.2, 11 September 2020, https://www.eosc-hub.eu/news/new-briefing-paper-cross-border-services

borders. The majority of research in Europe is funded nationally. As reported[52], it is challenging to coordinate such funding to meet the needs of international research collaboration, particularly multidisciplinary research collaboration because service providers are bound to their given national and/or community mandates. Funding sources are varied, complex and involve a large number of different rules, which contributes to suboptimal use of the combined Member States' investment in research resources. Consequently, there are many potential funding schemes and mechanisms that could fund the different components of the MVE but each comes with its own constraints and integrating them into a comprehensive funding plan will be a challenge requiring effort of an entrepreneurial nature to actively seek funding opportunities and secure them.

The demand for cross-border use of research resources clearly does exist and will continue to grow, notably to address the OECD Sustainable Development Goals[53]. The EOSC-hub and OCRE projects note that the main barriers to cross-border service provision are not technical but legal, financial, organisational and regulatory. These barriers include access policies and required levels of assurance; variations in Value Added Tax rules; data protection and intellectual property restrictions.

Research Infrastructures have circumvented such barriers by making use of in-kind (rather than in-cash) contributions that avoid many of these cross-border issues and are a common means for Member States (MS) and Associated Countries (AC) to contribute to such international undertakings.

As an example, ELIXIR[54] has a mixed funding model including in-kind and cash contributions coming from a number of mostly public sources. The ELIXIR Hub is funded through membership fees paid by Member countries, and much of this funding is then transferred back to Nodes to implement ELIXIR's five-year Scientific Programme. ELIXIR Nodes, which run the services that users access, are typically funded through national-level investments that support the national coordination within the Node. The development and operation of services is usually funded through national grants, sometimes through dedicated infrastructure grants but, more frequently, through competitive research grants where service development is one component.

In the ICT domain, PRACE offers another example of combining in-kind and in-cash contributions to deploy international HPC services based on agreements made between participating countries and sites. A fraction of the participating HPC sites' resources are made available to international users. Those countries that do not host HPC sites contribute to funding HPC operating expenses (and associate services) in order to access the resources. EuroHPC, as described in the section Connecting EOSC and EuroHPC above, brings together co-funding from hosting Member States and the EC for the next generation of HPC resources with formal agreements for access rights. Such agreements are necessary, especially for rivalrous shared services such as HPC involving compute, storage and network connectivity, to avoid abuse by users of shared resources.

So, while there are many examples of international research activities, they are bespoke arrangements that fall short of EOSC's objective to create a virtual environment to store, share and re-use research data across borders and disciplines. In effect, the digital single market for research does not exist and better alignment of national research funding

52 Supporting the Transformative Impact of Research Infrastructures on European Research, Report of the High-Level Expert Group to Assess the Progress of ESFRI and Other World Class Research Infrastructures Towards Implementation and Long-Term Sustainability, June 2020, doi: 10.2777/3423

53 http://www.oecd.org/dac/sustainable-development-goals.htm

54 Martin C, Smith A and ELIXIR Partners. ELIXIR's ELIXIR's Long-term sustainability plan (2019), https://doi.org/10.7490/f1000research.1117498.1

policies is necessary to bring it into existence and support the European Research Area[55]. EOSC could provide the opportunity to bring about such a policy change but it will require influence at a political level in the member states. The success of the recently created European COVID-19 research data platform provides an example of the advances that can be made if such a policy change happens. Taking inspiration from the Research Infrastructures and HPC examples cited above, EOSC should develop a funding model that combines in-kind and in-cash contributions based on agreements that align national research funding policies and leverage European-level funding schemes to ensure cross-border and cross-discipline research can flourish.

1.4.2 Business Models

As stated in the Final Report and Recommendations of the 2nd EOSC High Level Expert Group[56], the EOSC business model is a critical non-technical element that will determine the success of the EOSC vision. The business modelling activities of the EOSC core operational costs study suggested that the MVE - which includes EOSC-Core, federated data and EOSC-Exchange - is considered as an ecosystem to be sustained by a combination of platform business models. Platform business models create value by facilitating exchanges between two or more interdependent groups.

The study highlighted that two families of business models ('transactions' and 'learning') need to co-exist, potentially applied to different sides of the platform or targeting different clusters of roles and players, in order to sustain EOSC.

Transaction-based models are widely known and build on the perceived value in interactions between different entities. The platform facilitates transactions, reducing their costs and/or by enabling externalized innovation. Use cases analysed by the EOSC-hub project[57] highlighted that complex information needs to be accessed and exchanged before transactions between users and suppliers of research data, resources and services can be concluded.

The EOSC can add value by providing frictionless, easy access to data and related services so research communities can better connect with suppliers, users and funders by facilitating their interactions and offering matchmaking or brokering opportunities. EOSC can also promote a cross-fertilising multidisciplinary environment where investments can be efficiently leveraged and benefit from economies of scale.

Patronage/membership-based 'learning' business models promote the perceived value based on being part of a community and finding help and support or networking capabilities for their members. For example, in EOSC this could mean offering private dashboards to each research organisation through which they can track their consumption over longer periods, allowing them to negotiate better terms with the resource providers. Similarly, resource providers would benefit from continuous interactions with (potential) users, generating a private flow of data and insights to better tailor their future offers.

EOSC must allow the coexistence of different business models and their evolution over time. Different parts of EOSC can be based on different models. *Core* and federated data sections can benefit from a membership-based learning business model, partly funded by the EC and partly by the Member States. The mission of this part will be to provide support

---

55 A new ERA for Research and Innovation, European Commission, COM(2020) 628 final, Brussels 30.09.2020, https://ec.europa.eu/commission/presscorner/detail/en/ip_20_1749

56 Prompting an EOSC in practice, Final report and recommendations on the European Open Science Cloud (EOSC), 2018, doi:10.2777/112658

57 EOSC-hub Briefing Paper – Provision of Cross-Border Services, EOSC-hub project, version 2.1, 16 October 2020, https://www.eosc-hub.eu/publications

to the entire ecosystem to learn collectively how to allocate resources, how to manage cross-border research activities, and to provide shared cultural elements to spread best practices and seamless integration approaches. The *Exchange* section can benefit from a transactional model, pre-paid with monitoring usage of resources purpose for the commons, FAIR resources, open access resources available free of charge but with strong guiding and orienteering services. These models may dynamically change weights over time, building on lean thinking best practices and starting small.

1.4.3   Training and Consultancy

Qualified, experienced providers of expert technical support and consultancy are key to reducing friction in research transactions. Their input is required to successfully boost research collaboration and make more effective use of research assets. They can offer specialised support to help users to understand the implications of different choices, build demonstrators, assess technologies, and help arrange use of resources whilst protecting users from lock-in including supporting ease of movement between providers. Such expert technical support and consultancy roles have an established place in the publicly funded e-Infrastructures, notably OpenAIRE National Open Access Desks[58] and HPC Centres of Excellence[59] while the EOSC-hub Digital Innovation Hub[60] offers similar services to companies.

To make research more productive, EOSC needs to facilitate transactions between research users and suppliers by introducing standardised and optimised processes, regulations and agreements.

A key characteristic of a transaction is assigning a value to the resources being consumed but the data collected by two studies, EOSC core operational costs by AcrossLimits & Boundaryless[61] and the Landscape analysis by DCC, highlighted that the majority of resource providers and users are not aware of the costs involved. DCC notes that financial information is frequently missing from the country reports gathered as part of the landscape analysis and the interviews with service providers performed by AcrossLimits showed the majority of providers were unable to accurately estimate the costs of the individual services they offer. The EOSC core operational costs study showed this lack of information led to examples of under-estimating and over-estimating service costs.

The Virtual Access (VA) funding scheme as supported by the EC in Horizon 2020, has helped service providers to develop a better understanding of service costs. The study saw a higher-level of maturity in costing information provided by those service providers that had already determined VA unit costs. However, in most cases VA excludes capital investment from the calculation of the unit cost which implies service providers contribute with their own resources to the provision in order to comply with the co-financing and no-profit principles[62].

This suggests that the co-financing provided by the service provider should be recognised as an in-kind contribution and that, for private sector service providers, the services would

---

58 https://www.openaire.eu/what-is-the-openaire-network-noads

59 https://doi.org/10.5281/zenodo.3727821

60 https://eosc-dih.eu/

61 The Vivus Study , Ensuring that EOSC-Core & Minimum Viable EOSC are sustainable through a study on their costings, potential business models and funding schemes, October 2020, Angele Giuliano & Steve Robertshaw (AcrossLimits), Luca Ruggeri & Andrea Valeri (BoundaryLess)

62 DECISION authorising the use of unit costs for the actions involving virtual access

under the Research Infrastructures Part of the Horizon 2020 Framework Programme, accessed 29 August 2020, https://ec.europa.eu/research/participants/data/ref/h2020/other/legal/unit_costs/unit-costs_virtual-access_infra.pdf

need to be made available to publicly funded researchers at below market prices in order to comply with the no-profit principle.

1.4.4 MVE

The cost of operating *EOSC-Exchange* (but not the cost of operating the services it contains) will be included in the *EOSC-Core* funding model.

Based on the findings of the study on the EOSC core operational costs, it is estimated that the cost of operating *EOSC-Core* is approximately 7 million Euros per year. The level of confidence in the accuracy of this estimate is quite low due to a relatively superficial understanding of service costing by many of the publicly funded operators interviewed. It should also be noted that the costing currently does not include interoperable metadata and the operational support services which are, however, required.

The MVE is intended to assemble a set of services to benefit European scientists. In doing so, EOSC is assembling services that it does not own, and which are operated with very low overheads from funds typically allocated on an annual basis. The services are designed to serve prescribed communities of interest, which are often national communities. There is limited interest from the operators of these services to broaden access because there is no incentive to do so: funding is one problem together with the constraints of their local policies and regulations being another (for example, many service operators are only allowed to make a very small fraction of their capacity available to users beyond their original constituencies). Consequently, everything that needs to be done in order to achieve EOSC ambitions is (rightly) seen as a problem because the lack of overheads undermines operational flexibility. So, activities such as cross-border service delivery, VAT, procurement and micro-payments represent severe obstacles to service expansion and integration.

Funding for *EOSC-Core* in the long-term should be addressed by the participants engaging in EOSC[63] as a means of supporting open research for their communities. There are a number of formulae that have been proposed for the sharing of the costs of operating *EOSC-Core* and, as the Science|Business Network has concluded[64] a progressive, long-term approach to funding the EOSC is needed:

> *"Although it will ultimately need to break even, the initial priorities for the EOSC have to be driving participation and usage. Like most private businesses, the EOSC will probably need to operate at a loss (be subsidised) in its early years to ensure its proposition is appealing to both the data providers and the data users."*

For a new undertaking, such as the EOSC, significant initial financial engagement from the participants will be a challenge because participants are being asked to invest in a resource under development where the business case has yet to be demonstrated. To overcome this challenge, the financial risk for stakeholders is shared with the EC via the Co-Programmed Partnership (see Governance section) that sees a gradual shift of responsibility from the EC to the participants. Such a shift of responsibility is necessary because, as stated in the EC revised version of the Orientations towards the first Strategic Plan for Horizon Europe[65], partnerships also need to have a clear life-cycle approach, be time limited and include conditions for phasing out the Programme funding. As stated in the draft EOSC Partnership

---

63 Participants are expected to be drawn from a large group of organisations that goes beyond the membership of the EOSC Association

64 https://sciencebusiness.net/report/why-open-science-future-and-how-make-it-happen

65 Orientations towards the first Strategic Plan for Horizon Europe, European Commission, December 2019, https://ec.europa.eu/info/sites/info/files/research_and_innovation/strategy_on_research_and_innovation/documents/ec_rtd_orientations-he-strategic-plan_122019.pdf

Proposal, the need for an EOSC partnership will phase out when tools, systems and practices enabling Open Science will be widely available and deployed. Their uptake by research communities will then become the responsibility of local and national stakeholders in the Member States and Associated Countries and/or even at the level of individual organisations.

The EC will fund the *EOSC-Core* via the INFRAEOSC-03-2020 funding call for the period 2021-2023. During this period, more precise information about the operational costs of *EOSC-Core* must be collected and analysed. Taking into account information gathered about costs and usage, the EOSC Association should propose a cost-sharing mechanism and timetable for stakeholder financial contributions to be progressively increased over a number of years.

The scale and diversity of the services and resources to be federated implies that the operational and financial responsibility of federated services and data will remain with their existing operators and funders. The investment in federated services and resources by Member States needs to be measured and acknowledged as an in-kind contribution to the overall EOSC funding model.

Funds for developing, operating and maintaining the services included in *EOSC-Exchange* is principally the responsibility of the service providers that operate them. Services made available *via EOSC-Exchange* may be available free of charge or against payment but remain free at the point of use. *EOSC-Exchange* will establish trust between service providers and users by making use of the authentication and authorization services of *EOSC-Core* and providing service vetting features including reviews, profiles and certification schemes. Where a service is available against payment, charges will be transparent and visible, including *via* the portal. A transaction for the use of a service will be an agreement between the service provider and user or their sponsor (i.e. the operator of *EOSC-Exchange* will not be involved in the transaction*). To encourage the support for FAIR principles, the agencies and organisations (including European Regional Development Fund[66] and national programmes) funding research in countries participating in EOSC, should consider making a policy decision to accept the use of services in *EOSC-Exchange* as an eligible cost in data management plans and grant requests submitted by researchers.

1.4.5 Diverse funding sources

There are a range of private and public funding sources that could potentially contribute to establishing and perpetuating EOSC that have been identified by the EOSC core operational costs study, such as Digital Europe[67], Connecting Europe Facility (CEF2[68]) and European Institute of Innovation and Technology (EIT[69]) at the European level. Of particular interest could be the European Regional Development Fund (ERDF[70]) which, as part of Europe's cohesion funds, is frequently under-attributed across many of the Member States. Actions would be necessary at a political level to convince member state governments that EOSC related activities are a worthy target for such funding schemes, but their engagement could significantly contribute to the sustainability of EOSC.

1.4.6 Incentives for service providers and researchers

When Open Science and FAIR data become the defaults for research and the network effect brings its cumulative benefits, scientists will be more efficient in doing their research and

---

66 https://ec.europa.eu/regional_policy/en/funding/erdf/

67 https://ec.europa.eu/digital-single-market/en/europe-investing-digital-digital-europe-programme

68 https://ec.europa.eu/digital-single-market/en/connecting-europe-facility-cef2-digital

69 https://eit.europa.eu/

70 https://ec.europa.eu/regional_policy/en/funding/erdf/

society at large will increase its trust in science. At that point the return on investments in EOSC made by the stakeholders will have been achieved and the business case for them to fund its operation and expansion will have been demonstrated. To get EOSC started and reach that point, incentives as a temporary measure are necessary for service providers and researchers to engage in order to create the desired network effect.

It is recognised that the services provided by publicly funded organisations frequently have a mandate and a budget to serve a well-defined set of users that may be delimited by research discipline or geographical boundaries, and that broadening access to those services may generate additional costs. As an incentive to encourage service providers to participate in *EOSC-Exchange* and open up their services to all publicly funded researchers, the projects to be funded via calls such as INFRAEOSC-07-2020 will offer an EC-funded means, based on the Horizon 2020 Virtual Access[71] scheme, to compensate service providers for the additional operational costs they incur. Should a service provider request compensation (i.e. request Virtual Access funds) for providing a service *via EOSC-Exchange*, then it would be classed as a service available against payment.

As an incentive for commercial service providers to participate in *EOSC-Exchange* and offer services supporting the FAIR principles it is proposed that a centrally organised and aggregated procurement activity is undertaken by EOSC and funded by the EC. As an example of an incentive to commercial service providers, the OCRE[72] project has an EC funded budget of 9.5M€ to procure commodity commercial cloud services for the publicly funded research community. Its first tender[73] has attracted many providers, though issues with the handling of VAT payments persist and need to be addressed if such an approach is to be generalised.

Participation by service providers in the *EOSC-Exchange* incentive schemes described above will be subject to their commitment to participate in monitoring and reporting schemes intended to gauge usage and uptake of the services.

Access to the services included in the *EOSC-Exchange* Virtual Access and centrally organised procurement incentive schemes will be made available to public funded researchers via two distinct open call channels:

- Small-scale rapid access to resources for any publicly funded researcher (including long tail of science and citizen scientists) on request and subject to a basic eligibility check.

- Peer reviewed for larger-scale project access (similar to PRACE project access calls[74])

Under the Research Infrastructures Work Programme of Horizon 2020, the EC has provided support to actions that connect the ESFRI infrastructures to EOSC and promoted research data sharing by default. This has produced synergies that are acknowledged in an Independent Expert Report[75] which recommends a higher level of coherence in the funding

---

71 The Virtual Access (VA) instrument is provided by the European Commission to increase the sharing of research infrastructures and services that otherwise would not be available to international user groups. In VA, the services – also called "installations" – have to be made available 'free of charge at the point of use' for European or International researchers. VA access is open and free access to services through communication networks to resources needed for research, without selecting the researchers to whom access is provided.

72 https://www.ocre-project.eu/

73 https://www.ocre-project.eu/respond-tender

74 http://www.prace-ri.eu/call-announcements/

75 Supporting the Transformative Impact of Research Infrastructures on European Research - Report of the High-Level Expert Group to Assess the Progress of ESFRI and Other World Class Research Infrastructures Towards Implementation and Long-Term Sustainability, June 2020, doi: 10.2777/3423

from different chapters (RTD, CONNECT) of the Framework Programme, and across the three pillars of Horizon Europe.

The development of the ESFRI 2016 & 2018 roadmaps included an e-needs gathering activity that identified the ICT services of use to the proposed research infrastructures[76]. Estimating such needs is essential for planning the scope and capacity of resources to be federated via EOSC.

The National Science Foundation (NSF[77]) in the USA has taken the process of aligning funding programmes a step further by allowing researchers to request funds to access centrally procured cloud services to support their e-needs while submitting research grant proposals. NSF has funded CloudBank[78] as a 5 Million US$ pilot to enhance the research and education community's access to cloud computing resources through selected programs within the NSF Directorate for Computer and Information Science and Engineering (CISE). Grant applicants include a supplementary document that provides the cost details (which count toward the overall proposal limit), justification, and description of the cloud computing resources requested.

If such an approach were mapped onto EOSC, it would mean funding calls under Horizon Europe should allow applicants to request access to resources in EOSC-Exchange as part of their proposals. Such a funding model would allow researchers to plan their e-needs by consulting the EOSC catalogues of services and could help ensure the results of publicly funded research are made more widely available and accessible beyond the lifetime of the grants.

Such a model could also be extended to research grants funded via the European Research Council (ERC) and Marie Skłodowska-Curie actions as well as national and regional funding programmes.

1.4.7   Incentives for innovation

To encourage the development of innovative services supporting FAIR principles as well data stewardship and preservation across different phases of the research lifecycle, dedicated incentives schemes funded by the EC are foreseen that would use *EOSC-Exchange* as a distribution channel:

- Research and Innovation Action grants to develop services to be made available *via EOSC-Exchange*

- Pre-Commercial Procurement (PCP)/Public Procurement of Innovation (PPI) co-funding financial instrument for innovative services to be co-developed with the private sector, procured jointly by public authorities and commercialised *via EOSC-Exchange*

All such innovation incentives would require developments to adhere to Rules of Participation[79] resulting in production quality (Technology Readiness Levels 7-9) services to be included in *EOSC-Exchange* with associated training material and support.

---

76 Guide to e-Infrastructure Requirements for European Research Infrastructures, e-IRG, 1 March 2017, http://e-irg.eu/documents/10920/363494/2017-Supportdocument.pdf

77 https://www.nsf.gov/

78 https://www.cloudbank.org/

79 Rules of Participation, Version 0.5 (20 October 2020)

1.4.7.1   Innovation Grants

As an example of a cross-discipline research and innovation action providing incentives to Research Infrastructures (RIs) and the private sector, the ATTRACT[80] initiative brings together Europe's fundamental research and industrial communities to lead the next generation of detection and imaging technologies. Co-funded by the EC, ATTRACT has created a co-innovation ecosystem of 170 projects between fundamental research and industrial communities to develop breakthrough detection and imaging technologies for scientific and commercial uses. Operationally, it is based on the concept of "cascade funding" via an open call that attracted 1,211 proposals of which 170 were selected. Each of the 170 projects receive 100,000€ lump sum over one year to achieve a Technology Readiness Level 2-3 (TRL) and must produce a series of public deliverables. The projects are deep-tech initiatives covering some 15 domains, from fundamental physics to life sciences, environment and security and offer promising uses in fields such as imaging, cancer treatment, particle optics, sensor manufacturing, data streaming and storage as well as many more. Preliminary feedback shows that innovation actions have enabled 1 out of 3 projects to obtain additional private or public (e.g. national or regional) funding. It is envisioned to continue in a second phase with the objective of scaling up the most promising opportunities developed during the first phase. While the initiative is focussed on detection and imaging technologies, the same framework could be applied to other domains.

EOSC could offer a dissemination and commercialisation path for the outcomes of innovation actions such as ATTRACT and many of the projects have expressed their willingness to make further results openly available including data, published papers, technology application ideas etc. They could also benefit from other research assets made available via EOSC, such as data sets available within the same domain of their research, especially for example, related to bio-imaging. The networking capabilities apparent in the EOSC ecosystem, with many universities, SMEs, start-ups, corporations as well as research and technology organisations and RIs interacting via the same platform, offers a matchmaking or brokering opportunity for the formation of consortia that can respond to open calls operated as cascading grants.

1.4.7.2   Innovation Procurement Instruments

The use of innovation procurement instruments, such as PCP and PPI, is expected to increase in the next multiannual Framework Programme. The PCP instrument (90% co-funded by the EC in Horizon 2020) is used to steer the development of solutions towards concrete public sector needs, whilst comparing/validating alternative solution approaches from various providers.

The PPI instrument (35% co-funded by the EC in Horizon 2020) permits public procurers to act as launching customer / early adopter / first buyer of innovative commercial end-solutions newly arriving on the market. PCP and PPI actions can be applied sequentially, so that PPI can be used to procure the results of an earlier PCP at a larger scale.

The results of the completed PCP/PPI projects in the ICT domain[81] funded via FP7, CIP and Horizon 2020 show the instruments facilitate access to the procurement market for SMEs; bring research results from the university to the market; reduce the R&D risks for procurers and encourage commercialisation of the results by vendors; improve the quality and efficiency of public services; and contribute to growth and jobs in Europe.

---

80 https://attract-eu.com/

81 Innovation Procurement - The power of the public purse - EU funded projects in the ICT domain, 2019, https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=5443

In the ICT domain, innovation procurement instruments have been successfully used to develop HPC services for PRACE and have bolstered Member State engagement in EuroHPC[82]. In the context of EOSC, HNSciCloud[83] has developed innovative data intensive science services that are available via the EOSC catalogue[84] while ARCHIVER[85] is developing archival and preservation services for research data that cover the full research lifecycle.

The consortia contracted via PCP tenders frequently include universities and research performing organisations working with companies that are capable of commercialising the results. Such tender specifications can avoid vendor lock-in by requiring solutions that adhere to international standards and privilege open source software that avoid excessive licensing costs and where transparency contributes to building trust in security and data protection.

The attention to commercialisation and costing that is a key feature of the innovation procurement instruments could offer a channel for sustaining the results of the projects funded via calls including INFRAEOSC-07-2020 and INFRAEOSC-03-2020 as well as research and innovation actions such as ATTRACT.

1.4.8   Summary of funding schemes

As presented in the sections above, there are many potential funding schemes and sources that could contribute to the funding of the different components of EOSC MVE. Each funding scheme has its own advantages and constraints that need to be carefully examined and accommodated in order to form a coherent sustainability plan. A theme that is common across all of these schemes is that EOSC is going to need strong political support in the Member States and Associated Countries to be able to access possible funding streams. For example, if European Regional Development Funds (ERDF) is to become a possible source of funding for EOSC it will require local and national support across Europe to make this happen. The cross-border delivery of research services is constrained by a number of non-technical obstacles that inhibit EOSC from achieving its multidisciplinary and multinational objectives. So, here again EOSC stakeholders will need to persuade their national governments to allow EOSC participants to provide services to other Member States as long as these are part of the agreed spectrum of joint EOSC activities.

A key recommendation of the EOSC core operational costs study is to **establish an EOSC funding support team** dedicated to identifying and securing funds, that is made up of entrepreneurial individuals whose aim is to ensure that no stone is left unturned when it comes to sourcing funds for EOSC. This team would need to liaise with the governance representatives of the Member States and Associated Countries to garner financial support for the development and expansion of EOSC. Close collaboration with the networks of National Contact Points for funding programmes such as Horizon Europe could compliment the competences and reach of this funding support team.

Further recommendations identified by the EOSC core operational costs study include:

- **Raise awareness within the research communities of what is EOSC-Core and provide a clear mapping of the services it contains**
  The concept of federating core functionality is relatively recent and needs to be

82 Kick-starting EU cooperation: Joint steering of solution

development through Pre-Commercial Procurement (PCP), https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=66524

83 https://www.hnscicloud.eu/

84   T-Systems   https://marketplace.eosc-portal.eu/services?providers=102   ;   Exoscale   https://marketplace.eosc-portal.eu/services?providers=97 ; Sixsq https://marketplace.eosc-portal.eu/services?providers=100

85 https://www.archiver-project.eu/

disseminated across the communities, so it can be integrated into their planning activities.

- **Promote wider use of cost-based accounting**
  Service costing activities need to move beyond FTE based project accounting in order to establish more accurate estimates of costs for service operation and future development. EOSC needs to engage with financial officers at service operators as part of this process.

- **Virtual Access is a step in the right direction but needs improvement**
  The Horizon 2020 Virtual Access funding mechanism encourages service providers to move beyond current project-based accounting approaches but needs to evolve further if it is to be more widely used in the longer-term.

- **EOSC requires Service Provider Aggregation**
  Ultimately it should be possible for the EOSC Association to procure services directly from service providers, but the current service landscape is fragmented and relies on service provider aggregators (GEANT, EGI, EUDAT and OpenAIRE) to structure the offers of individual providers. The service provider aggregators hold important knowledge about a wide range of services that is essential to EOSC and needs to be channelled via projects funded via calls including INFRAEOSC-03-2020 and INFRAEOSC-07-2020.

- **Pre-Commercial Procurement**
  The EOSC Association could coordinate and procure resources by implementing the more suitable norms and tools provided by the EU Procurement Code. As an example, the PPP (Public Private Partnership) and the PCP (Pre-Commercial Procurement) may be suitable tools to deal with the high expectations, co-design approach and strong flexibility required by the research context.

- **Consider EOSC sustainability as a whole**
  Going beyond the focus on EOSC-Core is needed to assess how the funded EOSC related projects will be interacting with each other and to understand how these projects will contribute to the sustainability of EOSC. EOSC-Core is focusing on horizontal, federating services, which are typically a pure cost for platforms. They are fundamental because the entire operational model relies on those components and modules to be executed, but their marginality is leaning towards zero. So, a platform strategy should cover their cost with other sustainability streams, coming from the transactional or from the "learning" sustainability models.

- **EOSC as a "super" platform**
  EOSC can become a "super-platform", with the possibility of coordinating and connecting other existing platforms such as those operated by the service provider aggregators (GEANT, EGI, EUDAT and OpenAIRE) and the Research Infrastructures in a way that both they and EOSC would benefit by this strategy. By coordinating the existing platforms, EOSC would enhance important data flows that can be reused to provide better services and shared with the coordinated platforms and other third-parties. At the initial stage, the EOSC platform should relay, include and aggregate the already existing platforms and give them more chances to improve and scale across the national borders.

- **Enhancing the (business-like) skillset present in the EOSC community**
  There is an abundance of highly skilled and motivated people working in the EOSC community to ensure it is a success. However, they are struggling with some of the more business-like skills, such as accounting, market analysis and business planning. It would benefit the community and help to more efficiently achieve sustainable operations, if such skills could be engaged more readily in EOSC activities.

2   EXPANDING EOSC: ENGAGEMENT OF THE WIDER PUBLIC SECTOR AND PRIVATE SECTORS

The initial iteration of EOSC will focus on establishing the Minimum Viable EOSC (MVE) to address the needs of publicly funded researchers exploiting openly available data. Subsequent iterations will expand the EOSC to address use cases beyond openly available

FAIR data and engage a wider user base including the public sector and the private sector.

In order to successfully extend the EOSC knowledge ecosystem beyond the core research community, EOSC must demonstrate value and impact that is relevant and meaningful to the diverse groups belonging to broader public and private sectors.

There is potential to widen the circles of EOSC knowledge stakeholders in phases through existing strategic alliances and by means of progressive expansion of knowledge across all categories of stakeholders, starting from the inner circles of EU consortia, Public-Private Partnerships (PPPs), to sector-specific and citizen bodies, and further on to citizen engagement groups.

EOSC should aim to expand to include public and private stakeholders who form part of the wider EOSC knowledge ecosystem and thus enabling further excellence by the European research community. Through the involvement of the private sector in EOSC, research and the private sector can interact through the use of data and services marketplace. Industries could see their involvement by adding value services on research data of commercial relevance but may also consume scientific outputs. The value proposition EOSC offers to the private sector as a service provider/developer includes:

- a commercialisation channel to promote services via the service catalogue and marketplace;

- a simplified procurement channel for selling to public research sector organisations;

- a means of interacting with researchers to improve services.

Moreover, to the private sector as a consumer, the EOSC value proposition is to offer a mechanism for accessing:

- publicly funded open access/FAIR datasets;

- a range of services that can help them exploit open access/FAIR datasets;

- a means of interacting with the groups producing the datasets and services.

The *MVE* can be expanded with additional functionality and services dedicated to the requirements of end-users from the public sector[86] who are not involved in research activities but want to exploit open access to research data. For example, EOSC could support education use-cases that could potentially exploit the data and services made available *via* EOSC as Open Educational Resources (OER). EOSC can also offer assistance to the public sector in relation to the implementation of the Open Data Directive[87] into national law by Member States by July 2021. The scope of the Open Data Directive includes research data resulting from public funding and focuses on the economic aspects of the re-use of information. EOSC can assist with the publishing of dynamic data, the uptake of Application Programme Interfaces (APIs) and address the transparency requirements for

---

86 In this document the term 'public sector' refers to all bodies governed by public law as defined in Public procurement of services: Council Directive 92/50/EEC

87 Directive (EU) 2019/1024, https://eur-lex.europa.eu/legal-content/EN/TXT/?qid=1561563110433&uri=CELEX:32019L1024

public–private agreements involving public sector information, avoiding exclusive arrangements. The monitoring functions of EOSC could also help Member States identify high-value datasets associated with important benefits for the society and economy.

The *MVE* can also be expanded with additional functionality and services dedicated to the requirements of end-users from the private sector so that they can exploit the FAIR data and associated services for commercial gain without distorting market competition.

The European Investment Bank published a report[88] which included a section on the EOSC and found that

> *"The unique selling point (USP) of the EOSC is the magnitude of data in the context of the convergence of HPC, Big Data and machine learning."*

Feedback from the stakeholders interviewed in a recent study[89] suggest that EOSC's USP as the "Web of FAIR data" extends beyond the FAIRification of scientific research data. The study identified significant use cases in which data would be shared between external public and private sector organisations (especially in the use case of actively researching clinicians) that is currently not being shared because the data is not trusted. In addition to having the qualities of being Findable, Accessible, Interoperable and Reusable, the study proposes that data sets be created and used in a way that was verifiably JUST: Judicious, Unbiased, Safe and Transparent. Rather than describe the quality of the data, JUST refers to the actions of the data creator or user. By providing the FAIRification and JUST certification as a service, EOSC becomes a central hub in an exchange between external stakeholders that might otherwise not make use of the EOSC. As the Web of FAIR data, several industry stakeholders have suggested that EOSC services that allow for verification (or 'FAIRification') of industrial data sets would be welcome as a trustworthy standard to underpin industrial data exchanges. The means of verifying adherence to FAIR and JUST qualities would need to be introduced gradually in order to give the stakeholders time to adapt their datasets and services.

A category of external stakeholders broadly overlooked by the discrete categories of Industry, SME and Citizen Scientist is that of professional researchers and clinicians. Those who have emerged from the academic sector and now operate a professional practice collecting and applying big data, do so without the analytical support infrastructure of a research institution. By providing a bridge between external professional practitioners and academic research using FAIRification of data as a condition of participation, professionals can benefit from better data processing algorithms, while European researchers can make a significant contribution to e.g. patients of specialist clinicians to and society at large.

2.1 Use cases for the expansion of EOSC

The Industry Commons Foundation study on the Expansion of EOSC engagement to the wider public and private sectors identified use cases resulting from cross-domain, data-driven applications created in pan-European collaborations by research communities, citizen scientists, public sector organisations and industry, as well as new and emerging case studies from grass roots innovation communities, industry demonstrators and European projects. These use cases centred on the premise that expansion of EOSC beyond the public funded research sector must enable further excellence by the European research community in ways such as:

---

88 Financing the future of supercomputing How to increase investments in high performance computing in Europe, European Investment Bank, 2018, doi:10.2867/31460

89 Expanding EOSC: Engagement of the wider public sector and private sectors in EOSC, Michela Magas and Andrew Dubber, Industry Commons Foundation/ MTF Labs AB,9 October 2020

- stimulate novel research methodologies and support research excellence;

- enhance existing research practices through greater access to data based on FAIR data principles;

- encourage the development of novel Open Science research exchanges that allow interdisciplinary and international collaborations, and open up new categories and fields of knowledge;

- establish a mechanism for a technology transfer that is grounded in the research communities' ethos and principles;

- enable a marketplace for exchange of knowledge and datasets, established upon FAIR data implementation guided by the research communities' values.

The study has documented a total of 23 use cases which collectively suggest the expansion could monitor public health, the environment and the medical sector, foresee extreme climate weather and pandemic peaks, as well as help fine-tune energy use, transportation management and the setup of new materials. A selection of these use cases are described below:

Sentinel Hub: BlueDot Observatory

SMEs leveraging global monitoring of water bodies on a shoestring budget through API access. Managing water crises is one of the Sustainable Development Goals, and the significant decline in the available quality and quantity of fresh water has been ranked by the World Economic Forum's 2018 Global Risks report as one of the top ten most serious societal risks facing the world. This use case highlights the commercial and societal potential for European Open Research Data, but also the challenges faced by EOSC to act as an intermediary and an enabler in this context.

Open Media – European Broadcasting Union

Promotion of EU digital sovereignty and means of preserving and promoting cultural and historic value of European public media archives. The EBU use case describes a multiplier effect for news gathering and provision by providing instantaneous translation and targeted news aggregation and verification. The use case raises questions about the e-Infrastructure offering of data storage and processing at scale in competition with commercial providers for use in a public service media context.

COVID-19

The European readiness for future pandemics is of utmost importance and should be addressed to ensure the preparedness of infrastructure building on already existing frameworks, such as the COVID-19 Data Portal, for broader use such as the EOSC. The COVID-19 use case supports the widening of EOSC to the public and private sectors and help fast-track the global visibility of EOSC. Integration of molecular research data with sensitive patient and clinical data will ensure that patients benefit directly from the research supported by EOSC. Europe's industry, including SMEs, will access data and deposit data in the public domain. Cross-linking with socio-economic, societal response and other social science and humanities will promote an integrated understanding of European outbreak response and preparedness and demonstrate the value of FAIR data to society and public engagement during a global public health crisis.

Višnjan Observatory: Citizen science

As a member of the International Asteroid Warning Network (IAWN), Visnjan is amongst the top five observatories in the world in collecting more near-Earth object (NEO) measurements to determine if they are a threat to Earth. Without these so-called follow-up and confirmation measurements the majority of newly discovered asteroids that are daily discovered mainly from Hawaii, would get lost in a day or even in a matter of hours. Measurements are taken to ascertain if the discovered object is really there, calculate its trajectory and verify whether it is a potential threat. Višnjan is a member of Spaceguard Foundation, an association that supports the creation of a system to discover celestial bodies which could potentially be a threat to life on Earth. The use case demonstrates the impact and scientific gravitas of citizen science projects that exist outside academia and the potential for recognition and support through non-monetary incentivization mechanisms and acknowledgement.

PaNOSC

Contribution to the realisation of a data commons for Neutron and Photon science, providing Open Data services and tools for data storage, analysis and simulation, for the many scientists from existing and future disciplines using data from photon and neutron sources. This use case demonstrates the potential for innovative SME bridging organisations to translate large amounts of specialist scientific data to meet the needs of industry research and product development, and the potential for new markets to emerge based on European research.

Industry OntoCommons: Siemens Complex Equipment

Describing and analysing the digital twin of products/industrial assets in manufacturing and energy industry across their lifecycle from design to service based on IT systems. This use case demonstrates the importance and centrality of FAIR data in industry and the potential for EOSC to act as a Web of FAIR data in a context within which Industry is developing ontological interoperability.

BDVA: ICE Datacenter Gold i-Space

The ICE Datacenter Gold i-Space provides testing in a flexible full-scale datacenter - without large scale investment, with access to massive amounts of research data and with an on-call team of world-leading scientists who can contribute to an organisation's innovation activities. The Green Computing[90] use case, along with space data and other types of datasets, demonstrate the potential for Industry engagement with e-Infrastructures and a model for working with large research datasets for the private sector.

Human Rights Data: Cambridge Whisper

Collection and processing of highly sensitive and confidential data through interviews with refugees about their personal experiences of human rights abuses. This use case demonstrates the potential for unique tools that build upon the EOSC framework and portal, allowing for specific scenarios with software requirements that model best practice in the tools themselves.

Ocean Data

Navigating complex data sets and studies across a wide range of disciplines in the EU Oceans Mission in order to initiate agile and adaptive prototyping projects that give both citizens and industry the tools and autonomy to engage with and respond to a richer

---

90 A number of Sustainable Development Goals are addressed by the European Union Green deal with respect to computing. https://ec.europa.eu/digital-single-market/en/news/energy-efficient-cloud-computing-and-green-digital-services

understanding of seas and oceans. This use case demonstrates the potential for academic research to engage with citizen users in order to collaboratively address local challenges as well as those that affect industry and the environment.

Neurofeedback Patient Data

Clinicians collect and process large amounts of patient data from EEG Brainwave monitoring. There are significant challenges in storage and analysis of this data and enormous potential for anonymised data sharing that would reveal larger patterns and more nuanced understanding. This use case highlights the potential for EOSC to act as an intermediary Web of FAIR Data verification platform between non-academic professional researchers.

All the use cases have been analysed to derive an initial set of models and solutions for EOSC community engagement and long-term sustainability. The analysis of these use cases indicates that EOSC should act as the validating organisation for industrial FAIR data as well as for data by research communities. Beyond FAIR, the addition of JUST (Judicious, Unbiased, Safe and Transparent) which highlights accountability by a responsible researcher, has the potential to represent additional added value of EOSC to stakeholders. An additional important stakeholder group has been identified in professionals working with large valuable datasets (e.g. clinicians) who wish to be part of the EOSC marketplace.

From the analysis of the collected use cases, a set of recommendations has been created for EOSC financial sustainability implementation and expansion of EOSC with engagement of the wider public and private sectors:

- **Web of FAIR data**. For EOSC to have the greatest impact and reach to external stakeholders it must establish itself as the Web of FAIR data as its primary Unique Selling Point. Validation and interoperability of data in knowledge transfer and technology transfer are key to its centrality in the application (and collection) of research data from beyond the realms of academia.

- **INFRAEOSC-03 and INFRAEOSC-07 as environments for expansion**. The study has shown that the INFRAEOSC-03 and INFRAEOSC-07 funded projects should be used to initiate, implement, or prototype as appropriate, a series of actions recommended in this study that will enable the expansion of EOSC using the use-cases as concrete examples.

- **Alignment with EU marketplace initiatives**. Synergies with initiatives which are being developed in parallel, such as GAIA-X, EuroHPC, bloXberg, Industry Commons and the upcoming EIC marketplace have been highlighted by stakeholders and ought to be exploited, to save on duplication and speed up deployment of EOSC. These platforms require the same elements of marketplace infrastructure as EOSC and are all equally focused on data-driven use cases. If each of these marketplaces were to build its own decentralised system, not only would each be operating on different "railway tracks" but this would hinder ready adoption of common standards, useful data exchanges, interoperability, communications, knowledge and value exchange. The EU Common Market is transferring through digitalisation to a data-driven model, and the current version of the "EU Common Data Market", if left fragmented, risks ending up looking like a Bazaar.

2.2   Research Infrastructures engagement with industry

Engagement with industry as consumers of research output is happening today in many of the infrastructures to be federated via EOSC. As highlighted by ESFRI[91], usage by industry of RIs services and data is frequently by means of academic access in the framework of industry-university research contracts. EOSC could help RIs increase engagement with industry (particularly SMEs) by including training and exploitation of RI services and data whilst preserving the often-demanded competitive secrecy.

2.3   Common European Data Spaces

As suggested in the European strategy for data document[92], the experience gathered in deploying the EOSC for the research community will prepare the way for Common European data spaces in domains including industrial (manufacturing), finance and agriculture. The interconnection between EOSC and the Data and Information Access Services (DIAS) cloud-based platform that provides access to services based on the Copernicus earth observation data is expected to contribute to the establishment of common data spaces.

2.4   European Cloud Federation

The European Cloud Federation[93] and Member States' initiatives, such as GAIA-X, aim to create an open ecosystem of trusted tools, services, storage and computing capacities for companies and organisations throughout Europe. A recent announcement[94] demonstrates the political will expressed by 25 Member States to work together towards deploying resilient and competitive cloud offerings to provide the trustworthy data processing infrastructure and services that public administrations, businesses and citizens require in Europe. Synergies with such initiatives must be explored by the governance bodies currently being established for the next phase of EOSC (2021 onwards).

2.5   Digital Innovation Hubs

In the context of EOSC-hub, EGI leads the Digital Innovation Hub (DIH[95]) which acts as a virtual business accelerator supporting companies in accessing the digital technologies and services offered by the EOSC. It combines 4 main pillars to help companies become more competitive: Pilot design and co-design, technical access, training and support. The concepts and practice of DIHs and networks of hubs have evolved from the competence centres, being the core mechanisms to support digitisation, to the creation of structures where different public and private stakeholders work together to support SMEs as a Digital Innovation Hub. A network of DIH exists across Europe[96]. The EOSC DIH stimulates the innovation potential of research infrastructures, SMEs/industry, and other innovative actors. Establishing partnerships with SMEs allows for the increased exploitation potential of commercially viable research data and other existing e-Infrastructure services, and the provision of commercial services to researchers.

---

91 Supporting the Transformative Impact of Research Infrastructures on European Research, Report of the High-Level Expert Group to Assess the Progress of ESFRI and Other World Class Research Infrastructures Towards Implementation and Long-Term Sustainability, June 2020, doi: 10.2777/3423

92 A European strategy for data, 19.2.2020, COM(2020) 66 final, https://ec.europa.eu/info/sites/info/files/communication-european-strategy-data-19feb2020_en.pdf

93 https://ec.europa.eu/digital-single-market/en/cloud

94 Towards a next generation cloud for Europe, 16 October 2020, https://ec.europa.eu/digital-single-market/en/news/towards-next-generation-cloud-europe

95 https://eosc-dih.eu/

96 https://s3platform.jrc.ec.europa.eu/digital-innovation-hubs-tool

EOSC DIH has successfully run a series of pilots[97] that were selected as a result of the open call procedure. The pilots highlighted the importance of consultancy and support personnel to assist the companies, especially SMEs and the availability of high-quality production services.

The recently started EUHubs4Data[98] with an initial federation of 12 DIHs covering 10 countries and 12 different regions, aims to become the European reference for data driven innovation and experimentation, fostering collaboration between data driven initiatives in Europe, federating solutions in a global common catalogue of data services, and sharing data in a cross-border and cross-sector basis.

The EC foresees to support European Digital Innovation Hubs as part of the Digital Europe Programme[99] where co-investment from Member States is expected. If EOSC can make it simpler for RIs and e-infrastructures to participate in a network of European DIHs and better engage with industry, then this would represent a clear benefit to the EOSC stakeholders and the private sector.

As mentioned in the section [Training and Consultancy](#) above, HPC Centres of Excellence exist that support usage by the private sector and a fraction of PRACE resources (10%) is made available to companies[100]. The resources are free of charge for them, but they must agree to the PRACE Open Science rules of engagement. Moreover, any company may be a partner of academic proposals submitted to PRACE calls.

To summarise, EOSC can provide a societal challenge-driven environment for public and private sectors to co-design innovative data rich services and in turn increase Europe's technological autonomy and digital sovereignty in key enabling technologies and infrastructures for the data economy. Public-private cooperation enabled by EOSC can help address the low level of existing cooperation across Europe, where the EC estimated[101] that the percentage of innovative firms cooperating with knowledge institutes is only 15% and the share of public research financed by the private sector is only 7.2%. Through the ensemble of activities of the research and e-infrastructure communities, EOSC offers many opportunities to include public and private stakeholders as part of the wider EOSC knowledge ecosystem.

---

97 https://eosc-dih.eu/pilots/

98 https://euhubs4data.eu/

99 https://ec.europa.eu/digital-single-market/en/european-digital-innovation-hubs-digital-europe-programme-0

100 https://prace-ri.eu/prace-for-industry/industry-access/

101 A new ERA for Research and Innovation, European commission, COM(2020) 628 final, Brussels 30.09.2020, https://ec.europa.eu/commission/presscorner/detail/en/ip_20_1749

3 GOVERNANCE

The EC has demonstrated its support for EOSC during Horizon 2020 and its future role and engagement are dependent on the implementation of the next Framework Programme, Horizon Europe. To this end, the EOSC Governance Board and Executive Board have studied the suitability of European Partnerships as funding mechanisms for EOSC. European Partnerships are initiatives where the EU together with private and/or public partners commit to jointly support the development and implementation of a programme of research and innovation activities. Their added value lies in particular in bringing together a broad range of actors to work towards a common vision and translating it into concrete roadmaps and coordinated implementation of activities. European Partnerships need to have a clear life-cycle approach, be time limited and include conditions for phasing out the Programme funding.

3.1 Co-programmed Partnership

The EOSC Association was established on 29 July 2020 as a Belgium not-for-profit international association (AISBL). The EC intends to contribute to the funding of EOSC via Horizon Europe as a Co-programmed Partnership. One of the primary missions of the Association is to be able to enter into a Co-programmed Partnership with the EC for this purpose. However, the Association's purpose goes beyond this scope and timeframe.

The EOSC Partnership brings together all relevant stakeholders to co-design and deploy a 'Web of FAIR Data and Related Services for Science' and the EOSC Association is the focal point for this EOSC Partnership. The Association is a legal body, open to organisations from the public and private sectors that share the vision of EOSC and adhere to its values. An open and inclusive European Partnership will help ensure directionality (common vision and objectives) and additionality (complementary commitments and contributions at all levels). At inception, the EOSC Association is a relatively small-sized organisation that is expected to grow and evolve with EOSC.

A Partnership Board will be established as the main mechanism for dialogue in order to reach the objectives of the Partnership. In addition, a Strategy Committee representing Member States and countries associated with the Horizon Europe Framework Programme will be formed within the EOSC Partnership but outside the governance of the EOSC Association. Representatives from this Committee, together with delegates from the Commission and Association, will form the three parties in the Partnership Board of the EOSC Partnership. Additionally, the Strategy Committee will provide policy and strategic-level advice to the Association and overall Partnership from the perspective of the Member States and Associated Countries.

3.2 Vision and roadmap

The vision and roadmap of the EOSC will be developed by the Association in an open and transparent manner with the research and innovation stakeholders. One of the first tasks of the Association is to develop this Strategic Research and Innovation Agenda (SRIA), the guidelines that help develop the work programme for EOSC in Horizon Europe. Considered to be a living document, the first version of the SRIA was prepared by the current EOSC Executive Board and will be available by the end of 2020. It will accompany the Memorandum of Understanding (MoU) that will be signed between the EC and the EOSC Association to officially kick-off the EOSC Co-programmed Partnership. The signature of the MoU commits the Association to the specified objectives, key performance and impact indicators, outputs to be delivered, as well as the related financial and/or in-kind contributions of the members.

3.3   Participation in the Association

The internal organisation of the Association will integrate the views of all stakeholders. It should thereby be impartial in order to ensure a separation between oversight and operation and allow an objective assessment of the progress, uptake and effectiveness of its activities. The Association adheres to the following principles:

- Decentralization: the legal entity is as small as possible

- Transparency: the statutes, bylaws, the membership contract and any other organisation document are public

- Openness: any new member is able to join at any time, subject to compliance with simple rules of participation for new members

The Association encourages a broad spectrum of stakeholders to join EOSC, ensuring a balanced representation regarding types of infrastructural, organisational and sectoral members as well as geographic spread. This includes research data infrastructures, research performing and research funding organisations, researcher associations, and public and commercial service providers. Organisations based in EU Member States and countries associated with the most recent Framework Programme for research can join as full members, and organisations based elsewhere are able to participate as observers.

The feedback received while developing this document and its earlier versions highlight a divergence of objectives, priorities and interests across and within stakeholder groups concerning the benefits and costs of EOSC as well as the disruption it may cause to existing structures, decision making processes and funding models. The Association will have an important role to play in building trust amongst its members to resolve such divergence by offering a forum where the views of the stakeholders can be heard and taken into account so that the Association can confidently represent its members and speak with one voice.

An external Strategy Committee representing Member States and countries associated to the Horizon Europe Framework Programme will sit outside the Association. Its role will be to provide advice at policy and strategy levels.

The cost of running the EOSC Association[102], but not developing and operating EOSC, is covered by its membership fees which are approved by the General Assembly. While the EOSC Association's call for members has attracted an impressive number of candidates it is important that the membership fees do not become an obstacle to participation in EOSC. Ensuring that those organisations that cannot pay membership fees or are not legal entities can participate in the stakeholder forum will provide a channel for the voice of the broader research community.

The Association will coordinate the identification of needs for the development of EOSC and will provide input to the draft EC Horizon Europe work programme. Once the final work programme has been adopted by the EC, funding calls will be opened under well-defined topics to implement those elements of EOSC where there is a need for pan-European collaboration and funding. As a principle, the Association would not bid to Horizon Europe or other calls for proposals where this would represent a conflict of interest or competition with its members. Hans Craen, who led the team at Kellen Group that performed the study on Legal and strategic advice on EOSC legal entity membership and governance structure, noted during the EOSC Governance symposium, for the EOSC Association to be sustainable

---

102 The preliminary budget calculation foresees an initial running cost of ~500,000 euros/year and increasing to ~2 million euros/year. Final figures will be defined by the EOSC Association general assembly.

itself it must provide added value to its members. As described in the section Funding scheme above, the Association can also add value by actively seeking funding opportunities for its members to pursue EOSC activities.

3.4   Coordination of EOSC activities

During the course of the Partnership, the Association will coordinate EOSC related activities within its remit. It will focus on technical, communication and administrative roles. The technical role will essentially bring consensus and convergence in defining or contributing to the development and adoption of standards and good practices, within EOSC and globally.

The coherence of the programme and the synergies (internal and external) will be ensured by the Secretariat of the association. While being as lightweight as possible, the Secretariat will primarily focus on technical (coordination) and communication roles.

The Association should have a role in the definition of policies and the adoption of appropriate standards for the EOSC ecosystem. Similarly, the Association should also have a role in the monitoring and enforcement of adherence to defined policies and standards by its members and ensuring a risk management plan, as described in the Risk Management section below, is implemented.

If needed, the Association will exercise or outsource operational responsibilities such as managing specific services. This role will be limited in size and avoid conflicts of interest with its membership. When the resources necessary to operate a service become too significant, responsibility will be either transferred or hosted by another entity in a transparent and cost-effective manner.

Projects funded via actions of the EC Horizon Europe work programme will contribute to EOSC implementation strategy, for example, by delivering services for the *EOSC-Core* and *EOSC-Exchange*. Services funded by other channels, for example by national funders and private sector providers, are also expected to contribute to the EOSC.

The EOSCsecretariat.eu project organised project coordination events[103] throughout 2019 that brought together many of the projects contributing to EOSC. These events gave rise to the formation of several interest groups with participation from project representatives that contributed to a better coordination. Such events and coordination activities should be repeated under the new governance structure and could be organised by the INFRAEOSC-03-2020 funded project as one means of reviewing the progress of the collaboration agreements it has to establish with those projects funded under INFRAEOSC-07-2020.

3.5   Identifying and monitoring contributions to the partnership

As stated in the introduction, there needs to be political will not just towards the notion of open research, but also to fund the means to perform open research in practice. Governments have to be in a position to supply such funding, in accordance with a jointly agreed cost sharing model, and to coordinate their approach, policies and legislation. Here the Association can also play a role: It could for example leverage its Mandated Organisations[104] as a channel for communicating with Member States and Associated

---

103 Building EOSC through the H2020 projects current status and future directions, Belgium, 9-10 September 2019 https://www.eoscsecretariat.eu/events/building-eosc-through-h2020-projects-current-status-and-future-directions ; EOSC Coordination Day, Hungary, 28-29 November 2019, https://www.eoscsecretariat.eu/events/eosc-coordination-day

104 Mandated Organisations are defined in the EOSC Association's statutes https://www.eoscsecretariat.eu/sites/default/files/eosc_statutes.pdf

Countries, in order to develop a funding model that combines in-kind and in-cash contributions based on agreements that align national research funding policies and European-level funding schemes to ensure cross-border and cross-discipline research can flourish. The Association can also play a role in monitoring such contributions to the EOSC Partnership.

3.6   Communication

The promotional role of the Association will be based on multidirectional communication, actively supporting the users' engagement and feedback in all shapes and forms, as well as promoting EOSC results and success stories to convince new users of 'what's in it for them'. The Association will also, on behalf of its stakeholders, communicate with the EC and the society at large by speaking with one voice.

A schematic representation of the governance structure is shown in the figure below.
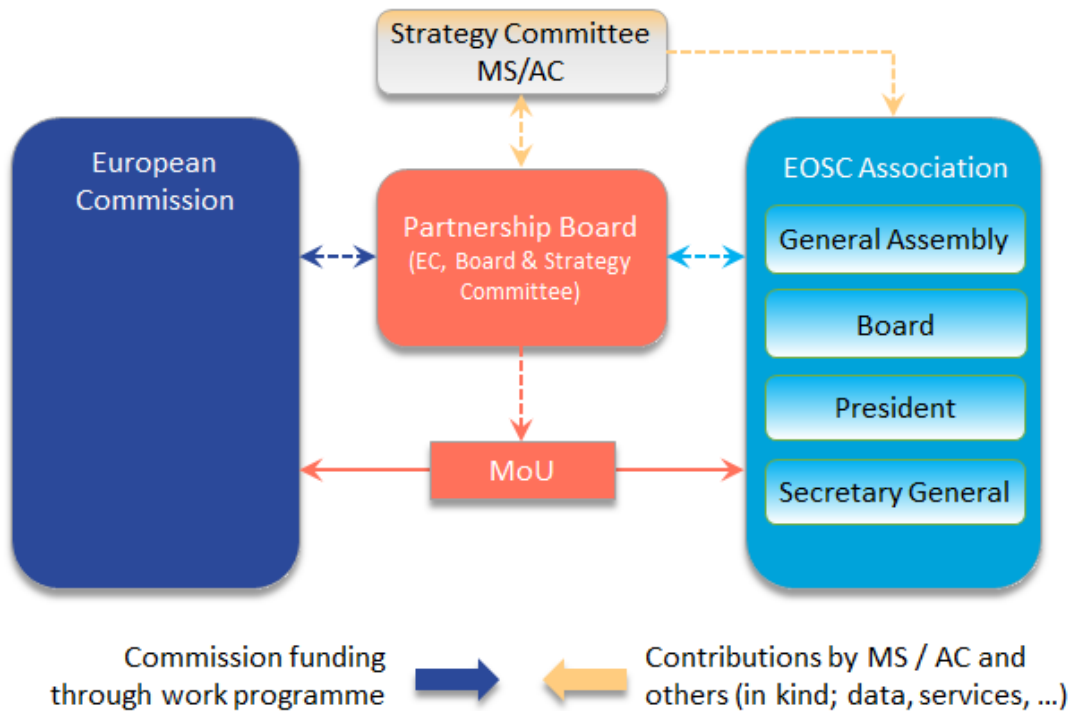
Figure 2: Schematic representation of EOSC governance structure

3.7   Risk Management

Complex organisations such as EOSC need to consider risk management to overcome crises and achieve their strategic objectives. Such Enterprise Risk Management (ERM) includes the activation of an organised risk management structure, as well as a pool of actions and processes, rules and responsibilities through which decisions are taken and implemented in terms of risks. ERM is a system of competences, organisational roles, policies, processes and models of analysis that allow an organisation's management to improve governance and control over its development path.

ERM is aimed at increasing the value of an organisation for the benefit of its stakeholders, supporting its objectives through the preparation of a methodological framework that allows a coherent and controlled performance of each future activity, the improvement of the decision-making process, planning, and creating priorities through a comprehensive and structured understanding of the activity itself. Risk management also contributes to a more effective use and allocation of capital and resources within the organisation, to the protection of the assets, to the corporate image, to the know-how of the organisation and the key people, as well as to the optimization of operating efficiency.

A targeted study[105] has been conducted by AON Hewitt in order to introduce a clear, and structured guidance on how to incorporate risk management into the governance of the EOSC Association. A risk model was developed defining five specific risk categories for the EOSC Association:

- **Operational**: is the prospect of loss resulting from inadequate or failed processes, procedures, systems, people, policies or technologies;

- **Reputational**: reputational risk refers to the potential for negative perception that the wide spectrum of stakeholders connected with EOSC may have. Reputation represents a successful indicator for the entire EOSC ecosystem.

- **Governance**: is any of various types of risk resulting from an inadequate and ineffective governance

- **Financial**: is any of various types of risk associated with financing, funding and economic sustainability in the long term;

- **Strategic**: strategic risk is associated with misalignment of the strategy to the mission and vision of EOSC, to the failure of including the strategy in the decision making and failure in deployment and/or execution of the strategy.

This has enabled the identification of 48 gaps in the risk governance with respect to best practices and highlighted that EOSC operates within a **multiple factor environment** with a high degree of complexity affecting the governance structure. The factors include the organisational model, political influences, multinational and cross-disciplinary usage. Risk management activities for EOSC to-date have been limited to individual project-based analysis and therefore fragmented meaning a clear and defined risk governance structure with assigned roles and responsibilities for risk management needs to be established.

This study sets out 32 recommendations to address these gaps and ensure the effectiveness of an EOSC risk governance. The main recommendations being as follows:

---

105 EOSC Risk Governance, Support the strengthening of the EOSC Risk Governance through the implementation of an effective Risk Management System, Final Report – September 30th, 2020

- **Launch a comprehensive plan** to address these gaps and define a risk governance framework and organisation to support the structuring and development process of EOSC itself;

- **Establish a governance structure for risk management** that is clear and well formalised with appointed roles and responsibilities across the organisational structure. It will be necessary to clarify the responsibilities for the different actors involved.

- **Map the skills and competences** required to perform an effective risk management at different levels of the organisational structure in order to consider all the fields of competence involved and set requirements on the composition of risk management bodies to **assure independence in decision-making**.

From an operational point of view, it is important for EOSC to **set-up an infrastructure and data security team** focused on the MVE with responsibility to:

- design a process that ensures the quality of the research data and data services;

- design, update and share cyber security, business continuity and disaster recovery policy;

- define a catalogue of potential risks (e.g. cyber-attacks, business interruption, damage to data, failure of systems or applications, etc);

- improve technical resilience of the MVE by:

o performing specific business impact analysis and identifying the most relevant business interruption risk causes;

o establishing and updating the business continuity management plan;

o preparing and testing the disaster recovery plan;

o defining a set of guidelines concerning resilience, business continuity and disaster recovery for service providers.

The identified gaps and recommendations are to be considered by the EOSC Association, the EOSC contributing projects and the EOSC Partnership overall in order to develop a comprehensive ERM.

Implementing these recommendations will significantly increase the value of EOSC and benefit its stakeholders by supporting its objectives and allow for a more effective use and allocation of resources. The ERM will also help to protect the assets, the corporate brand, the know-how of the key people, and optimize the operational efficiency.

During its study, AON has found a human capital very rich in multidisciplinary technical skills, sensitivity to governance issues, passion for the activities to be carried out and for the belief in EOSC itself. Moreover, the presence of all, or almost all, the essential pillars for the construction of effective risk management has been found.

4 TIMELINES

Implementing the *MVE* via a series of iterations described in this document will take the EOSC schedule beyond the end of 2020 as depicted in the graph in Figure 2.
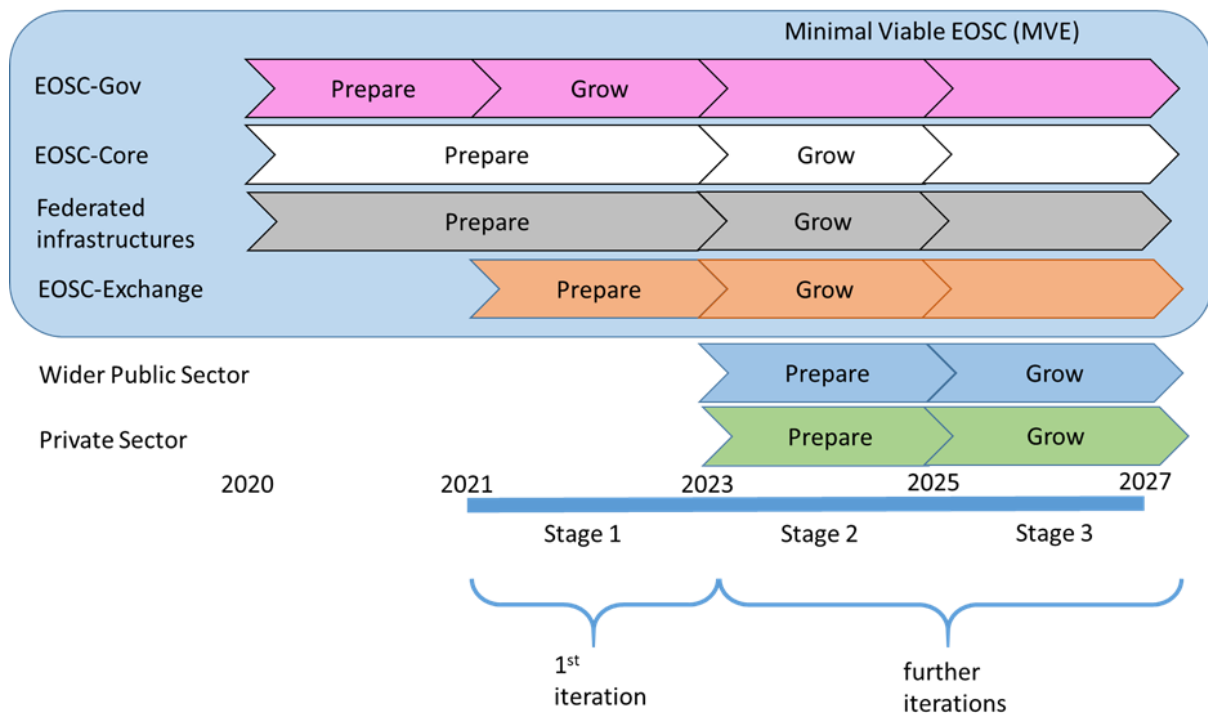

Figure 3: Schematic representation of timelines of EOSC iterations

Consequently, it is recommended to foresee a transition period of 3 years (2021-2023) to establish *MVE* building on projects to be funded via Horizon 2020 funding calls including INFRAEOSC-03-2020 and INFRAEOSC-07-2020, EOSC-related projects as well as contributions committed by the EOSC Association members. Collectively, the thematic clusters of Research Infrastructures and the e-infrastructures projects[106] have already set up the prototype system for the *MVE* and operate the EOSC Portal as an example of how to access such resources.

---

106 EOSC-related e-infrastructures projects (indicative): EOSC-hub, OpenAIRE Advance, eInfraCentral, EOSC-Enhance, FREYA

5 RECOMMENDATIONS

This section summarises the recommendations discussed in earlier sections of the document.

Recommendations on funding schemes and business models:

- **Establish an EOSC funding support team**
  A dedicated to identifying and securing funds, that is made up of entrepreneurial individuals whose aim is to ensure that no stone is left unturned when it comes to sourcing funds for EOSC. This team would need to liaise with the governance representatives of the Member States and Associated Countries to garner financial support for the development and expansion of EOSC and allow EOSC participants to provide services to other Member States as part of the agreed spectrum of joint EOSC activities. Close collaboration with the networks of National Contact Points for funding programmes such as Horizon Europe could compliment the competences and reach of this funding support team.

- **Raise awareness within the research communities of what EOSC-Core is and provide a clear mapping of the services it contains**
  The concept of federating core functionality is relatively recent and needs to be disseminated across the communities, so it can be integrated into their planning activities.

- **Promote wider use of cost-based accounting**
  Service costing activities need to move beyond FTE based project accounting in order to establish more accurate estimates of costs for service operation and future development. EOSC needs to engage with financial officers at service operators as part of this process.

- **Virtual Access is a step in the right direction but needs improvement**
  The Horizon 2020 virtual access funding mechanism encourages service providers to move beyond current project-based accounting approaches but needs to evolve further if it is to be more widely used in the longer-term.

- **EOSC requires Service Provider Aggregation**
  Ultimately it should be possible for the EOSC Association to procure services directly from service providers, but the current service landscape is fragmented and relies on service provider aggregators (GEANT, EGI, EUDAT and OpenAIRE) to structure the offers of individual providers. The service provider aggregators hold important knowledge about a wide range of services that is essential to EOSC and needs to be channelled via projects funded via calls including INFRAEOSC-03-2020 and INFRAEOSC-07-2020.

- **Long-term preservation of data**
  Digital preservation is not explicit in the context of EOSC and the roles, responsibilities and accountability for digital preservation are currently not clearly defined although the availability of long-term data preservation services represents an important added-value for EOSC. Clearer roles and responsibilities are needed and the assessment of capability as well as functions, salaries and funding streams for preservation should be explored.

- **Pre-commercial Procurement**
  The EOSC Association could coordinate and procure resources by implementing the more suitable norms and tools provided by the EU Procurement Code. As an example, the PPP (Public Private Partnership) and the PCP (Pre-Commercial Procurement) may

be suitable tools to deal with the high expectations, co-design approach and strong flexibility required by the Research context.

- **Consider EOSC sustainability as a whole**
  Going beyond the focus on *EOSC-Core* is needed to assess how the funded EOSC related projects will be interacting with each other and to understand how these projects will contribute to the sustainability of EOSC. *EOSC-Core* is focusing on horizontal, federating services, which are typically a pure cost for platforms. They are fundamental because the entire operational model relies on those components and modules to be executed, but their marginality is leaning towards zero. So, a platform strategy should cover their cost with other sustainability streams, coming from the transactional or from the "learning" sustainability models.

- **EOSC as a "super" platform**
  EOSC can become a "super-platform", with the possibility of coordinating and connecting other existing platforms such as those operated by the service provider aggregators (GEANT, EGI, EUDAT and OpenAIRE) and the Research Infrastructures in a way that both they and EOSC would benefit by this strategy. By coordinating the existing platforms, EOSC would enhance important data flows that can be reused to provide better services and shared with the coordinated platforms and other third-parties. At the initial stage, the EOSC platform should relay, include and aggregate the already existing platforms and give them more chances to improve and scale across the national borders.

- **Enhancing the (business-like) skillset present in the EOSC community**
  There is an abundance of highly skilled and motivated people working in the EOSC community to ensure it is a success. However, they are struggling with some of the more business-like skills, such as accounting, market analysis and business planning. It would benefit the community and help to more efficiently achieve sustainable operations, if such skills could be engaged more readily in EOSC activities.

Recommendations for expansion of EOSC with engagement of the wider public and private sectors:

- **Web of FAIR data**. For EOSC to have the greatest impact and reach to external stakeholders it must establish itself as the Web of FAIR data as its primary Unique Selling Point. Validation and interoperability of data in knowledge transfer and technology transfer are key to its centrality in the application (and collection) of research data from beyond the realms of academia.

- **INFRAEOSC-03 and INFRAEOSC-07 as environments for expansion**. The study has shown that the INFRAEOSC-03 and INFRAEOSC-07 funded projects should be used to initiate, implement, or prototype as appropriate, a series of actions recommended in this study that will enable the expansion of EOSC using the use-cases as concrete examples.

- **Alignment with EU marketplace initiatives**. Synergies with initiatives which are being developed in parallel, such as GAIA-X, EuroHPC, bloXberg, Industry Commons and the upcoming EIC marketplace have been highlighted by stakeholders and ought to be exploited, to save on duplication and speed up deployment of EOSC. These platforms require the same elements of marketplace infrastructure as EOSC and are all equally focused on data-driven use cases. If each of these marketplaces were to build its own decentralised system, not only would each be operating on different "railway tracks" but this would hinder ready adoption of common standards, useful data exchanges, interoperability, communications, knowledge and value exchange. The EU Common Market is transferring through digitalisation to a data-driven model, and the current

version of the "EU Common Data Market", if left fragmented, risks ending up looking like a Bazaar.

Recommendations on risk management:

- **Launch a comprehensive plan** to address the identified gaps and define a risk governance framework and organisation to support the structuring and development process of EOSC itself;

- **Establish a governance structure for risk management** that is clear and well formalised with appointed roles and responsibilities across the organisational structure. It will be necessary to clarify the responsibilities for the different actors involved.

- **Map the skills and competences** required to perform an effective risk management at different levels of the organisational structure in order to consider all the fields of competence involved and set requirements on the composition of risk management bodies to **assure independence in decision-making**.

- From an operational point of view, it is important for EOSC to **set-up an infrastructure and data security team** focused on the MVE.

This document explores possible means for sustaining the European Open Science Cloud beyond its initial phase which terminates at the end of 2020. This independent document builds on earlier strawman and tinman versions and the feedback received from the EOSC Executive Board, Governance Board as well as the European Commission (EC) and the stakeholder community on each version. It also takes into account the progress towards the EOSC goals as well as the outputs of commissioned studies.

It considers the financing model, legal vehicle, governance structure under the planned European Partnership with the EC as well as the regulatory and policy environment of the EOSC. It recommends beginning with a first iteration to establish a Minimum Viable EOSC (MVE) addressing the needs of publicly funded researchers exploiting openly available data. Subsequent iterations expand the EOSC to address usage beyond openly available FAIR data and engage a wider user base including the public sector and the private sector.

Research and Innovation policy

Publications Office
of the European Union