Financial Sustainability
Task Force

Towards Sustainable Funding Models for the European Open Science Cloud

Financial Sustainability Task Force Progress report
November 2022

Abstract

This progress report presents the main findings to date of the Financial Sustainability Task Force. The Task Force has scoped and defined the EOSC Exchange and Data Federation in the context of financial sustainability, identified barriers and boundary conditions, and performed initial development of financial scenarios for the EOSC Core, Exchange and Data Federation which address the period from 2027 onwards, after current EC Framework Programme seed funding for EOSC runs out.

EOSC Association AISBL

Rue du Luxembourg 3, BE-1000 Brussels, Belgium
+32 2 537 73 18 | info@eosc.eu | www.eosc.eu
Reg. number: 0755 723 931 | VAT number: BE0755 723 931

Members of the Financial Sustainability Task Force

Dale Robertson, EGI (Co-Chair)
Jan Meijer, Sikt–Norwegian Agency for Shared Services in Education and Research (Co-chair)

Bob Jones, EOSC-Association (board liaison)

Alberto Pérez Gómez, RedIRIS, Spain
Arnaud Roi, DARIAH Eric, Germany
Damien Lecarpentier, CSC, Finland
Françoise Hacque-Cosson, CEA, France
Ivana Ilijasic Versic, Elixir-Hub, UK
Jakub Szprot, University of Warsaw, Poland
Jan Hrušák, Czech Academy of sciences/ESFRI, Czech Republic
Jean-Pierre Vilotte, CNRS, Institut des Sciences de l'Univers (INSU), France
Jessica Klemeier, EMBL (alternate for Rupert Lück), Germany
Krzysztof Kurowski, Poznan Supercomputing and Networking Center, Poland
Lars Kaczmirek, University of Vienna, Austria
Luděk Matyska, CESNET, Czech Republic
Miguel Rey Mazón, Graz University of Technology, Austria
Natalie Haley, Instruct-ERIC, UK
Ornela De Giacomo, CERIC-ERIC, Italy
Patricia Mergen, Meise Botanic Garden, Belgium
Peter Kraker, Open Knowledge Maps, Austria
Ramaz Kvatadze, GRENA, Georgia
Rene Belsø, DeIC, Denmark
Rupert Lück, EMBL, Germany
Silvana Muscella, Trust-IT, Italy
Valter Nordh, NORDUnet, Denmark
Vanessa Proudman, SPARC Europe

Michel Schouppe, European Commission - DG-RTD (Observer)
Peter Szegedi, European Commission - DG-CNECT (Observer)

Acknowledgements

To the memory of our colleague Laura Perini, who passed away in August 2022.

The following Task Force members comprised the editorial group responsible for producing this document: Arnaud Roi, Dale Robertson (co-chair), Françoise Hacque-Cosson, Ivana Versic Ilijasic, Jan Meijer (co-chair), Jessica Klemeier (alternate to Rupert Lück), Lars Kaczmirek, Miguel Rey Mazón, Ornela De Giacomo, Patricia Mergen, Rene Belsø and Silvana Muscella.

EOSC Association AISBL
Rue du Luxembourg 3, BE-1000 Brussels, Belgium
+32 2 537 73 18 | info@eosc.eu | www.eosc.eu
Reg. number: 0755 723 931 | VAT number: BE0755 723 931

Table of contents

EOSC Association AISBL
Rue du Luxembourg 3, BE-1000 Brussels, Belgium
+32 2 537 73 18 | info@eosc.eu | www.eosc.eu
Reg. number: 0755 723 931 | VAT number: BE0755 723 931

EOSC Association AISBL
Rue du Luxembourg 3, BE-1000 Brussels, Belgium
+32 2 537 73 18 | info@eosc.eu | www.eosc.eu
Reg. number: 0755 723 931 | VAT number: BE0755 723 931

Executive summary

This progress report presents the main findings to date of the Financial Sustainability Task Force, in partial fulfilment of its Charter milestone of proposing "a first set of models [for the financial sustainability of EOSC], discussed and evaluated with relevant stakeholders"[1]. In working towards this milestone, the Task Force has scoped and defined the EOSC Exchange and Data Federation in the context of financial sustainability, identified barriers and boundary conditions, and performed initial development of financial scenarios for the EOSC Core, Exchange and Data Federation which address the period from 2027 onwards, after current EC Framework Programme seed funding for EOSC runs out.

The primary audience for this report is the EOSC Tripartite Governance, EOSC Association mandated organisations and Task Forces, and relevant experts such as EOSC project partners, ESFRI RIs and European e-Infrastructures, from whom feedback will be invited.

Context

EOSC is in a transition phase and is developing within a highly dynamic environment. An EC procurement call for the EOSC Core and elements of the Exchange is imminent, and EOSC is one of twelve common European Data Spaces[2] with which synergies need to be developed.

EOSC needs to overcome a number of challenges, including the highly fragmented research data landscape with widely varying levels of "FAIRness"; the complexity, variety and lack of overall coherence of the funding landscape when viewed on a European scale, resulting in suboptimal use of member states' investments in research resources; and greater ease of creating (or duplicating) resources than operating and sustaining existing ones in spite of their value to research communities. EOSC is in a sense taking a holistic approach in an ecosystem which has grown organically, raising political, organisational and financial challenges such as those entailed by cross-border delivery of data and services. The wider political and economic climate in Europe and globally is also challenging.

Against this backdrop, the Task Force has worked towards developing a sufficient understanding and description of what needs to be financed, so as to be able to identify several financial scenarios describing how this can be sustainably financed.

EOSC Core

The Core is the set of internal services that forms the federating foundation of the EOSC, facilitating its operation, in particular the large-scale brokering of research data and services. It consists of various components which constitute the technical underlay on which researcher-facing resources can be federated, integrated and used, as well as organisational processes required for the effective functioning of the EOSC. In the opinion of the Task Force, the Core should be viewed as a single element from a funding perspective.

[1] Available from https://www.eosc.eu/advisory-groups/financial-sustainability

[2] See https://digital-strategy.ec.europa.eu/en/library/staff-working-document-data-spaces

EOSC Association AISBL
Rue du Luxembourg 3, BE-1000 Brussels, Belgium
+32 2 537 73 18 | info@eosc.eu | www.eosc.eu
Reg. number: 0755 723 931 | VAT number: BE0755 723 931

Both the European Commission (EC) and member states (MS) have a strong interest in maintaining strategic control over the EOSC: ensuring inclusiveness, national relevance and synchronisation with other European and national initiatives.

The Task Force has concentrated on how to *fund* the EOSC Core rather than on how the Core services would be *operationally provided*, since its focus is sustainable funding. The Task Force's proposed scenario for the EOSC Core may be summarised as follows:

Funding

- The operational cost of the EOSC Core is to be funded jointly by the EC and the MS. This includes continuous incremental improvement aimed at operational relevance of current functionality, but not major investments in new functionality
- Countries associated with the Framework Programme for Research and Innovation should be able to contribute financially to the Core
- Financial contributions from third countries and others should be considered as additional sources of revenue (i.e. non-essential for operating the Core)
- In-kind contributions may be considered but do not confer the right to be exempted from the financial contribution.

Governance

- As the EOSC is an infrastructure first and foremost to serve the scientific communities, the research community must be strongly represented in the governance of the Core. Hence, the governance structure will comprise not only the MS and EC but also intergovernmental organisations such as ERICs and EIROs[3] which represent large RIs and bigger organised research communities
- A Scientific Advisory Committee (SAC), consisting primarily of representatives from those using EOSC, will review the Core services and provide recommendations to the Core governance on a regular basis based on their use, performance, usefulness, costs, etc. and compare to other service offerings available
- Involvement of countries associated with the current EC Framework Programme (ACs) as well as third countries in the governance, e.g. in an observer or advisory role may be considered
- Last but not least, the EOSC is a collective good for all researchers in Europe and offers access to the EOSC's data and resources irrespective of the status of their member states' contribution to financing the Core.

The Core should be funded by public money provided by the EC and MS. The presence of the EC as a co-funder ensures inclusiveness, and funding from MS will ensure they remain co-

[3] If the majority of the member states of an international organisation consists of EU countries - as is the case for ERICs and EIROs - it should be assumed that their financial contributions to the EOSC Core, similar to the national RIs, are covered by the respective financial contributions to EOSC from their member countries

EOSC Association AISBL
Rue du Luxembourg 3, BE-1000 Brussels, Belgium
+32 2 537 73 18 | info@eosc.eu | www.eosc.eu
Reg. number: 0755 723 931 | VAT number: BE0755 723 931

responsible for its deployment in their countries. Joint ownership goes hand-in-hand with joint funding of EOSC Core. The MS and EC will jointly decide the strategic direction of EOSC and align it with wider European vision and priorities; the presence of the EC helps to balance the representation of the MS; countries have a voice in governance and are co-responsible for EOSC deployment in their countries.

Strong representation of the research community is also required however, to ensure EOSC serves researchers' needs. Hence, in addition to MS, possible Associated Country (AC), and EC representation, a future Core governance will ideally also include large-scale strategic science activities such as the ERICs and EIROs. This could be achieved via a Scientific Advisory Committee. The involvement in EOSC of ACs, and that of third countries, should be further explored.

The EOSC Core should be financially sustained by the contributions of the MS and EC. In-kind provision of services for EOSC Core should not in any way interfere with the fees for EOSC to be sustainable, as it is important for adequate funding of the Core to be assured. Fees should be contributed by all funders, independently of their contribution in-kind (with or without cost reimbursement) to the EOSC Core.

All European researchers should be able to benefit from data and services in EOSC; individual researchers should not be punished and excluded, should their country decide not to pay the contribution.

EOSC Exchange

The Exchange is the EOSC's pan-European marketplace for research services. It enables the brokering of horizontal and thematic services between providers and researchers.
Analysis of the Exchange identified three distinct types of service provisioning, each supported by a different financial model:

- Centrally financed consumption of services - access to a certain amount of usage of selected services is made available to EOSC users, centrally financed by EOSC. This category is divided into two subsets:
    - A selective service portfolio of essential services (horizontal and thematic) which is 100% centrally funded
    - A small set of novel services which will receive temporary subsidies to initiate take-up in the research community
- Access to commercial services - procurement-compliant access to contracts with research-relevant commercial services
- Brokered not-for-profit services - community services brokered between the thousands of organisational participants in the EOSC with service transactions facilitated by the marketplace. This category constitutes the true marketplace of EOSC. It includes both horizontal and thematic services.

EOSC Association AISBL
Rue du Luxembourg 3, BE-1000 Brussels, Belgium
+32 2 537 73 18 | info@eosc.eu | www.eosc.eu
Reg. number: 0755 723 931 | VAT number: BE0755 723 931

The selective EOSC service portfolio would be a centrally financed set of horizontal, thematic and commercial services and resources targeting heterogeneous scientific domains and research communities. It would support the increased cross-border use of services arising from the interdisciplinary research and re-use of resources which EOSC will stimulate. The portfolio of temporarily subsidised novel services would follow the same financial scenario and logic as the selective portfolio, with a service's subsidy reducing over time.

The selection and management of these portfolios requires a researcher-driven governance structure for which a Scientific Advisory Committee (SAC) is proposed, with representatives from the different research communities. The amount of financing required will depend on which services are included in the portfolios and how much consumption is to be centrally financed.

Public sector entities cannot just buy commercial services, being required instead to follow a public procurement exercise. It adds value to offer in the context of EOSC a well-maintained portfolio of procurement-compliant agreements with research-relevant commercial service providers which is well integrated with the EOSC technical infrastructure. Again, strict portfolio management, e.g. by mandating an SAC, will be required.

The Task Force has begun considering the possible exchange facilitation and remuneration mechanisms which may be available for the brokered services. Options include a voucher/token model, direct payment, and a subscription or freemium model. These require further research, including whether they will always require a central EOSC entity to broker financial transactions or whether money can flow directly from the institution to the service provider.

The Task Force has identified three possible scenarios for the future of the EOSC Exchange:

- Establishing a pan-European single market for research services
- EOSC Exchange as a platform for window-shopping - an additional platform through which providers may display their services, involving no major change from today in how things are organised and funded
- Finding some middle ground between the first two options.

The first of these scenarios has been the focus of the Task Force's attention to date, whilst recognising that an incremental approach would be required; the other options have not been studied in detail yet. Consideration of how a pan-European single market for research services might be achieved led to a number of observations being made, as outlined below.

European digital research infrastructures operate in a complex environment. The majority of e-Infrastructures' funding is provided at the national level or below, although EC funding has a powerful influence. Member states financing European research infrastructures are willing to allocate both staff and national funds to operate joint European and cross-border e-Infrastructures which support thematic research communities, but they can be reluctant to co-finance generic European e-Infrastructures. For the EOSC Exchange to work and be sustained, funds need to be found to meet the marginal costs of the additional cross-border consumption

EOSC Association AISBL
Rue du Luxembourg 3, BE-1000 Brussels, Belgium
+32 2 537 73 18 | info@eosc.eu | www.eosc.eu
Reg. number: 0755 723 931 | VAT number: BE0755 723 931

of services arising through EOSC. There is a clear mismatch between MS and EC commitment and funding when it comes to the question of why, how and what to fund. Current conditions for funding can inadvertently stand in the way of funding e-Infrastructures.

This complex funding landscape is far from ideal for the long-term sustainability of EOSC. The current model in which national or EU funding is generally restricted to early-to-middle-stage innovation needs to change to keep pace with the rapid changes in the landscape. For EOSC to be sustained, the EC, Member States, Research Funding Organisations and Research Performing Organisations need to review their existing funding mechanisms and explore and experiment with new funding mechanisms to help sustain a growing and maturing research ecosystem for science and society.

The mandate of national service providers and institutions needs to change, to provide their services outside currently established boundaries, and/or against payment in the EOSC Exchange. There are political, legal, policy and cultural barriers to this in addition to the need to further research the transactions the marketplace could support. The legal-operational setup of EOSC from 2027 onward is also of relevance to the ownership of the EOSC marketplace and catalogue as well as other EOSC components.

Data Federation

The SRIA[4] states (p.107) that "EOSC will be primarily a federation of existing data and services where data remain in their current repositories and EOSC provides a means to make those data more broadly discoverable and interoperable", but how this will happen is still to be determined.

Regardless of the form of federation finally chosen, it is possible to envisage that the EOSC Data Federation (DF) should enable researchers to find and acquire data from multiple sources available at any of the levels of aggregation (local/institutional, national, thematic, European or international) through attribute-based discovery. A federation with these characteristics would lift the barriers to reusing data.

For the Data Federation, four use cases, provided by members of the Task Force, informed discussion. These related to Blue-Cloud, CESSDA ERIC, the Covid-19 Data Platform, and DiSSCo. This allowed several general observations to be made:

- Data and services must go together in the design of EOSC, despite the fact that their financial scenarios are different
- Federating data implies interoperability between five levels of aggregation - institutional, national/regional, European, Thematic and international - which can cause duplication and confusion about allocation of responsibilities

[4] https://www.eosc.eu/sria

EOSC Association AISBL
Rue du Luxembourg 3, BE-1000 Brussels, Belgium
+32 2 537 73 18 | info@eosc.eu | www.eosc.eu
Reg. number: 0755 723 931 | VAT number: BE0755 723 931

- EOSC must consider existing data federations and repositories on different levels and make them discoverable on its portal
- EOSC is being created within a wider landscape including Common European Data Spaces, GAIA-X, the Global Open Science Cloud and other initiatives
- EOSC has an incentive to establish itself in a global context
- A full assessment of the financial needs of establishing and running the EOSC Data Federation requires consideration of costs at European level but also at local and national levels
- The EOSC Data Federation must be created using existing infrastructures and thematic ecosystems without duplicating efforts.

Four potential architectural models for the EOSC Data Federation were also identified, although this is not necessarily an exhaustive list. These are:

1) **"Overlay":** to enable data discovery from any provider, EOSC could put in place an intermediary (software) layer, centrally managed and maintained by EOSC, that ensures data interoperability to the highest degree possible. This would come "on top of" (i.e. in addition to) any DF layers which already exist, e.g. in a specific scientific domain or geographical area.

2) **Metadata catalogue:** Instead of linking the data, a system to track metadata (i.e. not data itself, but their attributes or characteristics), organised into a catalogue, could be adopted by EOSC. This catalogue could then be used to achieve data interoperability, since having a coherent and consistent metadata classification scheme for data of very different nature allows to find and eventually use diverse types of data[5].

3) **Catalogue of data providers and platforms:** If a common interface to access the (meta)data cannot be implemented, an alternative solution would be to compile a list of all data providers and the platforms they use, indicating how to access single or combined data sources depending on the scientific discipline or geographic area.

4) **Natural evolution**: In the absence of the EOSC, existing data infrastructures will remain as the go-to sources for researchers. Convergence between them would continue to occur but it would be uncoordinated and slow, involving significant costs and implying the extension of the current suboptimal (inefficient) use of resources and siloed data landscape.

Conclusions drawn from comparison of the use cases identified categories of additional costs for the EOSC Data Federation. Several costs are implied at the EOSC level, although these are dependent on the complexity, or architecture, chosen for the EOSC Data Federation. In addition

[5] OpenAIRE's Research Graph (https://graph.openaire.eu/) is a possible candidate which could provide *part of* the basis for a centralised EOSC metadata catalogue, with FAIRsharing providing its (manually) "curated database, policy and standard metadata" as "an authoritative source".

EOSC Association AISBL
Rue du Luxembourg 3, BE-1000 Brussels, Belgium
+32 2 537 73 18 | info@eosc.eu | www.eosc.eu
Reg. number: 0755 723 931 | VAT number: BE0755 723 931

the costs of aggregation levels below EOSC - European/international, thematic, national, institutional - need to be considered. The categories can be summarised as the costs of

- Making data FAIR
- Making experiments reproducible
- Ensuring long-term access to data
- Federating data to EOSC.

Some of these costs appear at all levels (European, thematic, national, institutional), whereas others require a clear assignment of responsibilities, which has yet to be done (for example for data preservation). Issues and related costs which arise in relation to federation of data include the need to de-duplicate data (multiple entries of a single data set, for example); the costs and consequences of harmonisation - alignment of metadata schemas and associated access procedures, certification and validation of repositories and operation and maintenance costs, in particular what will happen to the current thematic data portals once a model for federating data into EOSC has been created; and the costs of legal and ethical issues including sensitive data, where we recommend creating a federated group of experts for the EOSC and also setting up a risk management contingency find.

EOSC Association AISBL
Rue du Luxembourg 3, BE-1000 Brussels, Belgium
+32 2 537 73 18 | info@eosc.eu | www.eosc.eu
Reg. number: 0755 723 931 | VAT number: BE0755 723 931

1. Introduction

This report was produced by the EOSC-Association's Financial Sustainability Task Force (TF-FinSust) and was formally adopted by the TF-FinSust on 10 November 2022.

The board of the EOSC Association established the Financial Sustainability Task Force on 31 August 2021 to:

> *"produce by 2023 a proposal for long term financial sustainability of the main building blocks of EOSC: EOSC-Core, EOSC-Exchange and the Federation of Data & Data Services as defined in the FAIR Lady report "Solutions for a Sustainable EOSC"*[6]

The Task Force's role is described in its Charter[7]. Its mission is to develop and validate scenarios for EOSC financial sustainability for the period of 2027 and beyond[8].

The Task Force began its work in November 2021, several months later than was envisaged when the task forces originally began being set up, and builds on the previous work done by the former Sustainability Working Group. The Task Force is also guided in its discussions by the EOSC Strategic Research and Innovation Agenda[9] (SRIA). The EOSC Partnership is now in place to implement the SRIA by designing specific actions to realise the priorities and objectives underlined in the Multi-Annual Roadmap (MAR)[10]. For the upcoming 2023 and 2024 period, the MAR calls for the need to identify and test feasible future financial sustainability models for the time after Horizon Europe as one of the key priorities[11].

The Task Force's charter includes in its key outputs, by Q3 of 2022, a "first set of models, discussed and evaluated with relevant stakeholders". That is an ambitious challenge - significantly hampered by ongoing lack of clear definitions for and scope of the Exchange and Data Federation components of EOSC - which has not yet been entirely accomplished, although a lot of progress has been made. This report thus serves as an intermediate Task Force milestone and a progress report: it outlines initial findings and a first set of (directions of) potential financial scenarios for the funding of the three EOSC components which could be adopted beyond the lifetime of current Horizon Europe funding mechanisms, i.e. from 2027 onwards. Despite the financial scenarios covering a period several years in the future, time is of the essence to put the necessary conditions in place to be able to act on the scenarios, e.g. budget lines, legal entities etc. Due to this pressure, the Task Force has not yet addressed in

---

[6] *Solutions for a sustainable EOSC: A FAIR Lady (olim Iron Lady) report from the EOSC Sustainability Working Group*, https://op.europa.eu/en/publication-detail/-/publication/581d82a4-2ed6-11eb-b27b-01aa75ed71a1

[7] Available for download from https://www.eosc.eu/advisory-groups/financial-sustainability

[8] The financing for the period 2023-2026 is assumed to be covered by existing funding, including the EC procurement expected to be launched in October 2022 for the EOSC Core and elements of the Exchange. In this report the funding of the Core and the Exchange have been considered separately and we believe their funding models should be independent of one another.

[9] https://www.eosc.eu/sria

[10] https://eosc.eu/multi-annual-roadmap-mar-consultation

[11] https://www.eosc.eu/news/mar-development-seen-driver-he-work-programme, p. 19

EOSC Association AISBL
Rue du Luxembourg 3, BE-1000 Brussels, Belgium
+32 2 537 73 18 | info@eosc.eu | www.eosc.eu
Reg. number: 0755 723 931 | VAT number: BE0755 723 931

any depth involvement in EOSC of the private sector, citizen science or involvement outside Europe[12].

Given that EOSC Exchange and Data Federation haven't been defined in sufficient detail to develop scenarios for their long-term financial sustainability, for these two building blocks the Task Force has worked towards developing a sufficient understanding and description of **what** needs to be financed, to be able to identify several financial scenarios describing **how** this can be financed, as well as challenges and boundary conditions around them. The EOSC-Core already benefited from a clearer definition, which permitted discussion to focus on developing a proposal for its financial sustainability. Ideally the TF also addresses the level of costs for the three EOSC components, as relative size and absolute cost. At the moment this information is not available hence not in this report.

In line with previous documents, such as the *FAIR Lady Report*, the MAR envisions different funding models for each component: "the Core will be supported in a common manner by [national and EU] public funds, while the large variety of services in the Exchange will be resourced using different mechanisms" (MAR, p.16 and 28). The Task Force recognises this likely diversity, and also that different governance structures may be required for different elements of EOSC.

EOSC needs to overcome a number of challenges. Some of the most prominent amongst them include:

- the highly fragmented landscape of research data and the proliferation of repositories at all levels (local, institutional, thematic European and international), with widely varying levels of "FAIRness". Some repositories constitute disconnected silos where data is largely unfindable, thus inaccessible, and definitely not interoperable. This clearly hampers data reuse, knowledge circulation and, more importantly, it significantly reduces the impact science has on society in the broadest sense. EOSC aims to establish data sharing practices according to the FAIR principles[13] for data from any discipline and level of aggregation, which implies interoperability between all of them if they are to be discoverable through EOSC.
- fragmentation happens however at a more fundamental level in that, as identified in the SRIA (p. 112), the majority of research in Europe is funded nationally according to independent national priorities, with very individualised activities; moreover, within each country, there exist complex and varied funding sources which follow a large number of different rules. From the point of view of the European research landscape

[12] We recognise, however, the significance of the expansion of the user base and note the finding of the "Expanding EOSC Study: engagement of the wider public and private sectors in EOSC", relating to EOSC's potential as a basis for non-academic knowledge exchange. See https://zenodo.org/record/4463437#.Yz_QlnZBwQ9 Publication date 30 October 2020 ICF Industry Commons Foundation, funded under EOSCSecretariat.eu.

[13] For an introduction on the FAIR principles, Wilkinson, M., Dumontier, M., Aalbersberg, I. *et al.* The FAIR Guiding Principles for scientific data management and stewardship. *Sci Data* **3**, 160018 (2016). https://doi.org/10.1038/sdata.2016.18

EOSC Association AISBL
Rue du Luxembourg 3, BE-1000 Brussels, Belgium
+32 2 537 73 18 | info@eosc.eu | www.eosc.eu
Reg. number: 0755 723 931 | VAT number: BE0755 723 931

as a whole, this results in a suboptimal use of the member states' investment in research resources[14].

- Project-based grant funding is focussed on delivering innovative results, often making it easier to create new (or duplicate) services and resources than to operate and sustain existing ones despite their value to research communities. This threatens investments already made and creates instability in the research service market and therefore has negative implications on research.
- Following these points, the research ecosystem is one which has grown organically and rapidly. EOSC is in a sense a holistic approach which is at odds with this, raising political, organisational and financial challenges, e.g. those entailed by cross border delivery of data and services.
- More generally, the current political and economic climate in Europe and globally is a challenging one in which to implement EOSC. Wider influencing factors include levels of government debt and the increasing cost of borrowing, high energy costs (affecting the cost of IT services amongst other consequences), budget pressures on universities, and increased levels of nationalism.

The primary audience for this report is the EOSC Tripartite Governance, EOSC Association mandated organisations and Task Forces, and relevant experts such as EOSC project partners, ESFRI RIs and European e-Infrastructures. A good understanding of the EOSC and its current developments is assumed.

The Task Force will invite feedback from key stakeholders based on this progress report. Despite not yet having fully developed financial scenarios for the EOSC, at this point we would like to gather feedback on our findings so far. The remainder of the Task Force mandate period will be spent, with support from the EOSC Focus project, developing more detailed financial scenarios and accompanying recommendations.

The report is structured as follows: after the introductory chapter, the Task Force provides a more precise scope and definition of each of the three main EOSC building blocks. The three following chapters outline financial scenarios for EOSC Core, EOSC Exchange, and the EOSC Data Federation. Finally, the report provides an outlook towards the Task Force's next steps.

It should be noted here that EOSC is in a transition phase: EOSC's first tripartite governance structure for the co-programmed European Partnership has been in place for less than two years, with numerous projects still contributing to building and developing EOSC. In addition, thirteen other EOSC Association Task Forces are progressing their work in parallel to the Financial Sustainability Task Force, an EC procurement call is due to be launched shortly for the EOSC Core and elements of the Exchange, the results of an EC-commissioned study to provide a full characterisation of the "European Research Data Landscape"[15] are due to be

[14] This is discussed in EOSC-hub deliverable D2.5, available from https://www.eosc-hub.eu/deliverable/d25-final-governance-and-sustainability-implementation-roadmap. See recommendations in section 4.2.

[15] https://research-and-innovation.ec.europa.eu/news/all-research-and-innovation-news/ec-kicked-study-will-provide-full-characterisation-european-research-data-landscape-2021-06-22_en

EOSC Association AISBL
Rue du Luxembourg 3, BE-1000 Brussels, Belgium
+32 2 537 73 18 | info@eosc.eu | www.eosc.eu
Reg. number: 0755 723 931 | VAT number: BE0755 723 931

published soon after publication of the present progress report, and in addition EOSC is one of twelve common European data spaces[16] with which synergies need to be developed. This means that the Task Force operates in a highly dynamic environment and to progress its work certain assumptions had to be made which will be revisited in the coming months.

[16] See https://digital-strategy.ec.europa.eu/en/library/staff-working-document-data-spaces

EOSC Association AISBL
Rue du Luxembourg 3, BE-1000 Brussels, Belgium
+32 2 537 73 18 | info@eosc.eu | www.eosc.eu
Reg. number: 0755 723 931 | VAT number: BE0755 723 931

2. Scope of the EOSC main components

As stated in the SRIA, "*EOSC is envisaged as a federation of infrastructures, forming a Web of FAIR Digital Objects and Related Services for Science*"[17]. It has the ambition "*to provide European researchers, innovators, companies and citizens with a federated and open multi-disciplinary environment where they can publish, find and re-use data, tools and services for research, innovation and educational purposes*"[18].

In line with its charter, the Task Force focussed on the main building blocks of EOSC: EOSC-Core, EOSC-Exchange and the Federation of Data & Data Services. The EOSC Future project[19], building on the work of the EOSC Architecture Working Group[20] has developed a detailed EOSC architecture diagram (see figure 1). The Interoperability Framework and Support Activities, particularly training, were not the focus of the Task Force's discussions but many of the costs arising from them would be incurred as part of preparing data and data services for federation, which is discussed in Chapter 5.

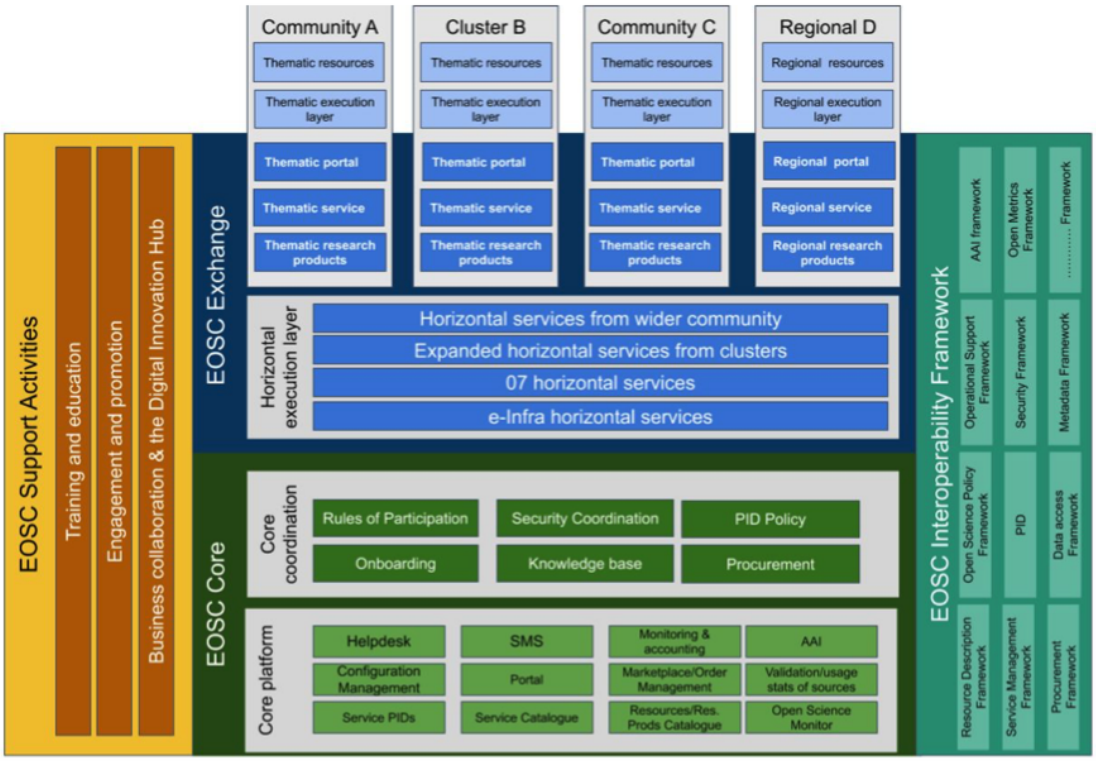


Figure 1: EOSC Future detailed architecture diagram

[17] p. 56 "Federation" is to be understood here as the adoption of a "federated architecture" of "loosely coupled" or decentralised "cooperating components", interconnected in a way "that allows interoperability and information sharing", see e.g. https://en.wikipedia.org/wiki/Federated_architecture.

[18] As indicated in https://eosc-portal.eu/about/eosc

[19] https://eoscfuture.eu/about/roadmap/

[20] https://www.eoscsecretariat.eu/working-groups/architecture-working-group

EOSC Association AISBL
Rue du Luxembourg 3, BE-1000 Brussels, Belgium
+32 2 537 73 18 | info@eosc.eu | www.eosc.eu
Reg. number: 0755 723 931 | VAT number: BE0755 723 931

Following the idea to determine **what** needs to be financed in order to be able to address the question of **how** to sustainably finance it and ideally **how much** it is anticipated it will cost, we present in this section a scoped view of the main building blocks of EOSC.

2.1 Core

The Core is the set of internal services that forms the federating foundation of the EOSC, facilitating its operation, in particular the large-scale brokering of research data and services[21]. The Core will consist of various components which constitute the technical underlay on which researcher-facing resources can be federated, integrated and used. The Core, as it is currently being developed, contains for example a service catalogue, AAI, monitoring, accounting, order management, helpdesk, etc. The Core also includes organisational processes required for the effective functioning of the EOSC: onboarding, rules of participation, security coordination, interoperability framework etc.[22] From a financing perspective, the Core should however be viewed as one single element with a minimal, defined, mostly stable set of services[23], without which the EOSC cannot exist as an operational reality.

2.2 Exchange

The EOSC Exchange is currently insufficiently defined to be able to identify what needs to be financed and how this can be financed. The FAIR Lady report envisages the Exchange as a "digital marketplace" (p11), which the SRIA states should be built "on the EOSC Core to ensure that a rich set of services (common and thematic), exploiting FAIR data and encouraging its reuse, are available to publicly funded researchers" (p. 68). From the Task Force's discussions it has become clear there are many different opinions on what the Exchange could and should be, and few definitions. Despite these differences, there is a common understanding that the Exchange has to facilitate cross-border service consumption, contribute to better services emerging through competition, and provide a fertile ground for niche services.

Services offered through the Exchange are usually separated in two categories; horizontal services providing generic services and resources which target heterogeneous scientific domains and research communities, and thematic (subject specific) services that target users from a specific science, community and/or regional domain. In our analysis, horizontal and thematic services have been treated in the same way although the funding of research infrastructures' services currently tends to differ from that of e-Infrastructures.

[21] EOSC participation is assumed to be large-scale: consumption and production of data and services by millions of researchers through thousands of organisations in a very dynamic environment. Large-scale federations take time to build and impose longevity requirements on the federation infrastructure.

[22] A detailed description of the EOSC Core is proposed in EOSC Future deliverable D2.5a, *Inventory of Core Functions and Inclusion Criteria*, available from https://cordis.europa.eu/project/id/101017536/results.

[23] The notion that the Core as the hub of the federation should be viewed as one coherent entity from the point of view of the federation participants does not mean it is not dependent on other infrastructures, such as nationally deployed identity federations enabling seamless pan-European login.

EOSC Association AISBL
Rue du Luxembourg 3, BE-1000 Brussels, Belgium
+32 2 537 73 18 | info@eosc.eu | www.eosc.eu
Reg. number: 0755 723 931 | VAT number: BE0755 723 931

Discussion in the Task Force led to the following scoping of the Exchange for the purpose of considering its financial sustainability:

The Exchange is the EOSC's pan-European marketplace for research services. It enables the brokering of horizontal and thematic services between providers and researchers. The Exchange encompasses a very heterogeneous set of services which makes a "one size fits all" funding model an inadequate solution. This report explores a scenario in which the Exchange is divided into three groups of services from a financial perspective:

- centrally financed services, formed of two sub-categories:
    - a small set of services flagged as essential for a broad range of research communities, which would have the same financial model as the Core but with an independent governance driven by the research communities
    - emerging services, which could temporarily and/or partially benefit from seed money until they achieve maturity;
- commercial services provided by the private sector,
- Brokered not-for-profit services, which would represent the largest subset of the Exchange and for which a limited number of standard remuneration mechanisms should be facilitated by the EOSC, including subscription, earmarked grants, vouchers, or others.

These categories will be further detailed in Chapter 4.

2.3 Data Federation

Access to community-endorsed reliable data and other research products stored in repositories anywhere in Europe is essential for EOSC to become useful. As already mentioned in the Introduction, the current situation in Europe allows this only in certain scientific areas where community-endorsed data standards and the required infrastructure has been implemented[24]. In other disciplines it is currently very difficult for researchers outside a given institution, region, or country to find and use data that do not "lie" in their immediate vicinity[25]. The SRIA states (p.107) that "EOSC will be primarily a federation of existing data and services where data remain in their current repositories and EOSC provides a means to make those data more broadly discoverable and interoperable", but how this will happen is still to be determined.

Regardless of the form of federation finally chosen, it is possible to envisage that the EOSC Data Federation (DF) should enable researchers to find and acquire data from multiple sources available at any of the levels of aggregation (local/institutional, national, thematic, European

[24] An example of this is CESSDA. See Appendix C.

[25] The existence of repositories in various research disciplines, like e.g. at EMBL (for life-sciences), COPERNICUS or Pangea (for earth and environmental sciences), to name just a few, is known to the Task Force. A complete assessment of the landscape is outside the scope of this report, although we understand a study providing a full characterisation of the European research data landscape is due to be published shortly - see https://research-and-innovation.ec.europa.eu/news/all-research-and-innovation-news/ec-kicked-study-will-provide-full-characterisation-european-research-data-landscape-2021-06-22_en.

EOSC Association AISBL
Rue du Luxembourg 3, BE-1000 Brussels, Belgium
+32 2 537 73 18 | info@eosc.eu | www.eosc.eu
Reg. number: 0755 723 931 | VAT number: BE0755 723 931

or international) through attribute-based discovery. A federation with these characteristics would lift the barriers to reusing data.

From a financing perspective, the Task Force considers the additional costs of establishing a possible federating structure sitting above the other levels of aggregation. It differentiates these from the costs of making data FAIR, ensuring long-term access, and enabling federation of data into EOSC on the local/institutional, national, thematic, European or international levels[26].

[26] The EOSC Association Task Force on Long-Term Data Preservation also, as its name suggests, studies issues of relevance to this. It is due to publish a report around the same time as the present progress report.

EOSC Association AISBL
Rue du Luxembourg 3, BE-1000 Brussels, Belgium
+32 2 537 73 18 | info@eosc.eu | www.eosc.eu
Reg. number: 0755 723 931 | VAT number: BE0755 723 931

3. Towards financial scenarios for the Core

The Core, the central element of the MVE (minimum viable EOSC)[27], will be funded until 2026 through the EOSC Procurement Action 2022-2023 run by the EC[28]. The Task Force's mission is, however, to explore how to finance the operation and further developments of the Core beyond that date.

Being a cornerstone of the whole EOSC, the Core should be under public governance, sustained through public funding i.e. the European Commission (EC), the Member States (MS), and the countries associated with the EC Framework Programme (AC). The crucial issue is the distribution of the financial commitment between those stakeholders and the management of the funding.

Discussion of appropriate funding models for the EOSC Core naturally gave rise to consideration of the requirements of a legal entity to handle the funding. Questions of governance of the EOSC Core were also closely related to this, and have been included in the Task Force's analysis although governance is beyond the scope of the Task Force's charter.

Instead of analysing the characteristics of existing legal instruments, the Task Force took the approach of identifying governance and funding requirements to guarantee a functional and sustainable EOSC (beyond funding itself, that is), which have been collected in Appendix A. This analysis may be used to identify a suitable legal instrument to support sustainable funding of the EOSC Core.

Funding

- The operational cost of the EOSC Core is to be funded jointly by the EC and the MS. This includes continuous incremental improvement aimed at operational relevance of current functionality, but not major investments in new functionality
- Countries associated with the Framework Programme for Research and Innovation should be able to contribute financially to the Core
- Financial contributions from third countries and others should be considered as additional sources of revenue (i.e. non-essential for operating the Core)
- In-kind contributions may be considered but do not confer the right to be exempted from the financial contribution.

[27] The minimal viable EOSC, or minimal viable product, is a term first used in the report of the 2nd EOSC High Level Expert Group, defined as "a product with just enough features to satisfy early customers, and to provide feedback for future product development". See https://op.europa.eu/en/publication-detail/-/publication/5253a1af-ee10-11e8-b690-01aa75ed71a1/language-en/format-PDF/source-80622260. The MVE was elaborated by the EOSC Architecture Working Group in its "View on the Minimum Viable EOSC", available from https://www.eoscsecretariat.eu/working-groups/architecture-working-group.

[28] Initial information on this procurement is available at the Tenders Electronic Daily website, https://ted.europa.eu/udl?uri=TED:NOTICE:234660-2022:TEXT:EN:HTML.

EOSC Association AISBL
Rue du Luxembourg 3, BE-1000 Brussels, Belgium
+32 2 537 73 18 | info@eosc.eu | www.eosc.eu
Reg. number: 0755 723 931 | VAT number: BE0755 723 931

Governance

- As the EOSC is an infrastructure first and foremost to serve the scientific communities, the research community must be strongly represented in the governance of the Core. Hence, the governance structure will comprise not only the MS and EC but also intergovernmental organisations such as ERICs and EIROs[29] which represent large RIs and bigger organised research communities
- A Scientific Advisory Committee (SAC), consisting primarily of representatives from those using EOSC, will review the Core services and provide recommendations to the Core governance on a regular basis based on their use, performance, usefulness, costs, etc. and compare to other service offerings available
- Involvement of countries associated with the current EC Framework Programme (ACs) as well as third countries in the governance, e.g. in an observer or advisory role may be considered
- Last but not least, the EOSC is a collective good for all researchers in Europe and offers access to the EOSC's data and resources irrespective of the status of their member states' contribution to financing the Core.

These proposals, and the Task Force's analysis which led to them, are explained below. Other questions connected with our proposals, e.g. on governance-related issues, are also included for future consideration, but we feel it is outside the Task Force's remit to provide comprehensive conclusions on all governance-related aspects.

Who should fund the EOSC Core?

The Core should be funded by public money provided by the EC and MS. There are several advantages to co-funding: the presence of the EC as a co-funder ensures inclusiveness, and funding from MS will ensure they remain co-responsible for its deployment in their countries[30].

The member states' contributions should be agreed once the legal instrument is selected. There is some solid experience in these funding models from ERICs and intergovernmental organisations (e.g. EMBL) which could be used to effectively achieve a consensus amongst funders.

[29] If the majority of the member states of an international organisation consists of EU countries - as is the case for ERICs and EIROs - it should be assumed that their financial contributions to the EOSC Core, similar to the national RIs, are covered by the respective financial contributions to EOSC from their member countries

[30] As of 26 October 2022, 24 member states have committed to Action 1 of the ERA (European Research Area) Policy Agenda "Enable Open Science, including through the European Open Science Cloud (EOSC)". Our proposal is consistent with this commitment.

EOSC Association AISBL
Rue du Luxembourg 3, BE-1000 Brussels, Belgium
+32 2 537 73 18 | info@eosc.eu | www.eosc.eu
Reg. number: 0755 723 931 | VAT number: BE0755 723 931

Who should govern the EOSC Core?

In the view of the Task Force, joint ownership goes hand-in-hand with joint funding of EOSC Core. There are several advantages in co-ownership and co-funding: the MS and EC jointly decide the strategic direction of EOSC, commit to it, and operate in synergy with other European initiatives. The presence of the EC helps to balance the representation of MS, including those with less resources to contribute to the EOSC. It also ensures the strategy of EOSC is in line with European vision and priorities. Similarly, funding from MS gives countries a voice in governance and makes them co-responsible for EOSC deployment in their countries.

However, the EOSC Core governance needs to include a strong representation of the research community too, to ensure EOSC continues to serve researchers' needs. Hence, in addition to MS, possible AC, and EC representation, a future Core governance will ideally also include large-scale strategic science activities such as the ERICs and EIROs, representing large RIs and big well-established research communities. This could be achieved via a Scientific Advisory Committee (SAC) reporting to the Core governance and consisting primarily of a broad representation from those using EOSC, including the long tail of science[31]. It would be the SAC's responsibility to review the Core services and provide regular recommendations to the Core governance on their usage, performance, costs, etc and to compare them to other service offers available.

The participation of international organisations in the governance of the EOSC Core deserves special treatment, at least for those with a majority of EU member states in their membership, as is the case for ERICs and EIROs. Here, it can be assumed that the contribution of an organisation to the EOSC Core would come from the respective national financial contributions to the EC's Framework Programme (which already includes contributions to EOSC), in a similar way as happens with national research infrastructures. International organisations could be exempted from contributing additional funding, to prevent MS paying twice for the same thing, but will still have a voice in the EOSC Core governance.

What is the role of Associated Countries and third countries?

The involvement of countries associated with the current EU Framework Programme in the funding and governance of EOSC, as well as that of third countries, should be further explored, considering the role they play in the European research landscape.

Our proposal does not preclude the existence of additional funding streams besides direct contribution from MS and (possibly) AC, for example by third countries. This remains to be further explored at a future stage (not necessarily by the Task Force).

[31] The expression "long-tail of science" refers to the fact that scientific communities follow a long-tailed distribution, with a few large ones and a large number of small ones.

EOSC Association AISBL
Rue du Luxembourg 3, BE-1000 Brussels, Belgium
+32 2 537 73 18 | info@eosc.eu | www.eosc.eu
Reg. number: 0755 723 931 | VAT number: BE0755 723 931

How to ensure a sustainable EOSC Core?

In the proposed financial scenario, the EOSC Core should be financially sustained by the contributions of the MS and EC. The Task Force has concentrated on the question of how to *fund* the EOSC Core, rather than on how the Core services would actually be *operationally provided*, since the Task Force's focus is sustainable financing. We include here however some observations relating to the operational sustainability of the Core.

It is unclear how the Core will transition from a short-term procurement to a long-term, sustainable infrastructure, since for example providers of services which form part of the EOSC Core may decide to discontinue their service, and some (or all) Core services may not represent a critical mass to discourage them from doing so. The Task Force would also like to draw attention to the fact that some public entities cannot participate in them as this is precluded by EU competition rules; therefore, limiting the procurement to tenders may preclude some public entities from applying. If some of the Core services would benefit from in-kind provisioning by a public entity, then this would need to be taken into account and regulated by a different type of procurement call (e.g. in-kind against payment). The outcomes of the EC procurement of the Core are not available at the time of writing this progress report but the Task Force will take them into account, once available, in its further work.

We believe however that in-kind provision of services for EOSC Core should not in any way interfere with the fees for EOSC to be sustainable, as it is important for adequate funding of the Core to be assured. Fees should be contributed by all funders, independently of their contribution in-kind (with or without cost reimbursement) to the EOSC Core. The reimbursement of the costs of Core services provided in-kind should come from the fees that all funders pay, according to previously agreed rules.

A core principle: access for all

All European researchers should be able to benefit from data and services in EOSC. Although the contributions by MS and the EC (and eventually third countries associated to Horizon Europe and future Framework Programmes) should provide for 100% of the operation and maintenance costs of the Core, individual researchers should not be punished and excluded, should their country decide not to pay the contribution. The Task Force recognises that the MS may wish to install a mechanism (for example removal of voting rights in the governance) to prevent a misuse of the all-inclusive model needed for EOSC Core services, where the majority use the Core while just a few contribute financially, but emphasises the importance of access for all users.

EOSC Association AISBL
Rue du Luxembourg 3, BE-1000 Brussels, Belgium
+32 2 537 73 18 | info@eosc.eu | www.eosc.eu
Reg. number: 0755 723 931 | VAT number: BE0755 723 931

4. Towards financial scenarios for the Exchange

In our initial analysis we have identified different options for the EOSC Exchange, including some considerations about their specific financial requirements:

1) **Establishing a single market for research services:** As defined in section 2.2, the Exchange shall provide the EOSC with a pan-European marketplace for research services. This marketplace will give researchers the freedom of choice to use that service which best fits their needs without being limited to their respective national, institutional or discipline-specific service provider.
2) **EOSC Exchange as a platform for window-shopping:** this would involve no major change from today in how things are organised and funded. EOSC Exchange will be an additional platform through which providers may display their services and thus increase their visibility beyond the research community in which they were established. Current challenges (outlined in chapter 1 of this document and further in this section) will persist.
3) **Finding some middle ground:** As will be pointed out again later in this report, data and services always have to go hand in hand in order to provide true added value. If the Exchange becomes a platform for window-shopping used to advertise services without further added value, the expected increased visibility for services will eventually leave most of the issues named in the introduction unsolved. Therefore, if the establishment of a single market for research services is deemed unfeasible - a point on which we wish to solicit feedback - the Task Force will explore further options on how to tackle some of the challenges identified.

In the following, we describe option 1, “Establishing a single-market for research services”, in more detail, outlining what this will look like as well as identifying challenges. Please note that the Task Force does not foresee such a change to happen overnight but as an incremental approach that will eventually lead to the described scenario. There has not been time to study the other options in detail yet although many of the challenges and conditions discussed below are relevant to all three options. The Task Force intends to perform further study of some of the key aspects relating to funding of the Exchange services in its future work.

Data services are currently replicated throughout the entire data life cycle across countries and institutions alike in a suboptimal and costly process. In a single market, this would gradually be replaced by a more efficient provision of higher-quality research services for the whole of the European Research Area which would allow best-of-breed services to be available to all European researchers regardless of where the service is provided from. Such a marketplace would afford service providers a much larger potential audience to cover service operation costs, enabling commodity services produced by the public sector research community to scale larger than currently possible, allowing niche services to reach a sustainable scale, and painting a target to aim at for projects establishing new services.

EOSC Association AISBL
Rue du Luxembourg 3, BE-1000 Brussels, Belgium
+32 2 537 73 18 | info@eosc.eu | www.eosc.eu
Reg. number: 0755 723 931 | VAT number: BE0755 723 931

A dynamic marketplace would mean that neither production nor usage of research services is hindered by borders. This would be more compatible with the inherently cross-border nature of research and cross-disciplinary research as well as contribute to the ERA and the challenges Europe is facing (e.g. health, green and digital).

Looking at the EOSC Exchange from a financial sustainability perspective, the Task Force identified three distinct types of service provisioning, each supported by a different financial model:

1) **Centrally financed consumption of services:** access to a certain amount of usage of selected services is made available to EOSC users, centrally financed by EOSC. This category is divided into two subsets:
    - **a selective service portfolio of essential services (horizontal and thematic)** which is 100% centrally funded
    - a small set of services, e.g. novel services, which will receive (temporary) subsidies to initiate take-up in the research community (**temporarily subsidised services**).
2) **Access to commercial services:** procurement-compliant access to contracts with research-relevant commercial services. This is put in a separate category because the gateway to entry is different from that for the other types of service provisioning described here; simply adhering to the RoP invites commercial suppliers to use the Exchange as a marketing brochure, whereas a ready-to-sign contract giving procurement-compliant access to the service makes it easier to use these services.
3) **Brokered not-for-profit services:** the potentially largest and most dynamic category consisting of community services brokered between the thousands of organisational participants in the EOSC, with service transactions facilitated by the marketplace as cost recovery mechanisms since no central funding will be provided through EOSC. This category includes both horizontal and thematic services.

4.1 Centrally financed services

4.1.1 Selective EOSC service portfolio

A certain amount of service usage of a subset of horizontal, thematic and commercial services and resources that target heterogeneous scientific domains and research communities, deemed essential for conducting research, will be grouped in a centrally financed portfolio. This portfolio will follow the same financial model as the EOSC Core i.e. it will be financed by a co-funding scheme from the European Union and Member States/Associated Countries.

The centrally financed portfolio will ensure the availability of certain standard services to all EOSC users in the entire ERA, regardless of national and institutional choices made, and thus help bridge the digital divide. To guarantee that the funded services match the ever-changing needs of researchers, a rigorous portfolio management shall be applied, with regular inclusion of new services and phasing-out and exclusion of outdated or unused services.

EOSC Association AISBL
Rue du Luxembourg 3, BE-1000 Brussels, Belgium
+32 2 537 73 18 | info@eosc.eu | www.eosc.eu
Reg. number: 0755 723 931 | VAT number: BE0755 723 931

Management of the centrally financed services portfolio requires a governance structure that will be responsible for setting the conditions for the selection of services, and for checking whether a service meets the EOSC RoP as well as the additional quality requirements which will apply to this service subset. Service selection needs to be driven by researchers' needs; the Task Force would therefore stress the importance of research communities having a strong voice in the portfolio governance. This will ensure the inclusion of bottom-up input useful to create momentum for EOSC.

In practice, this could be implemented in the form of a Scientific Advisory Committee (SAC) with representatives from the different research communities, who periodically review the pre-financed service portfolio and provide recommendations on the inclusion, phasing-out and exclusion of services. Having such a Committee in place is expected to lead to a high acceptance and usage rate among researchers as an additional advantage. The successful alignment processes among the different ESFRI cluster projects can serve as a basis for this committee. A question to be further explored is whether the SAC for the Core and the SAC for the collective portfolio of centrally-financed services could and should be the same or whether they should operate independently.

The amount of financing required will depend on which services are included in the portfolio and how much consumption is to be centrally financed. The Task Force will endeavour to investigate these questions further in its future work.

4.1.2 Temporarily subsidised uptake of new services

At any point in time there will be novel services whose uptake by researchers it would be desirable to stimulate. The Task Force proposes to create a dedicated facility for this, effectively a sub-category of the centrally financed services and following the same financial scenario and logic.

The subsidy may be a sliding scale (10-100%) of service consumption cost, to be reduced over time. Subsidised consumption could target specific areas where services beneficial for research outcome are insufficiently supported or adopted, e.g. geographies or disciplines.

As with the other centrally financed services, an SAC should be responsible for the management of the portfolio of temporarily subsidised services, assessing whether a (novel) service, e.g. resulting from an ending EC-funded project, is mature enough or has enough potential to receive temporary subsidy.

4.2 Access to commercial services

Commercial services are part of the researcher's toolbox, hence access to these should be provided through the EOSC. But what does that mean? Experience from GN4-3 and earlier GÉANT projects, OCRE and the EOSC-Future projects shows that allowing any and all commercial services to register on the marketplace following the gating process of the Rules of Participation will only lead to a large number of services on the marketplace which EOSC

EOSC Association AISBL
Rue du Luxembourg 3, BE-1000 Brussels, Belgium
+32 2 537 73 18 | info@eosc.eu | www.eosc.eu
Reg. number: 0755 723 931 | VAT number: BE0755 723 931

users cannot use[32]. Public sector entities cannot just buy commercial services, being required instead to follow a public procurement exercise. This is why access to commercial services requires a separate category in the EOSC Exchange.

It adds value to offer in the context of EOSC a well-maintained portfolio of procurement-compliant agreements with research-relevant commercial service providers which is well integrated with the EOSC technical infrastructure[33].

Institutions can inform their researchers that using commercial services is unproblematic as long as they are part of EOSC's portfolio of agreements and the researcher has arranged for the funding to pay for them. Researchers can inform their institution that the agreement is "safe" by virtue of being part of EOSC's portfolio of agreements.

A central portfolio of agreements in an EOSC context provides additional benefits:

- The trend is for commercial service providers to grow bigger and bigger, and the relative share of R&E of total provider revenue is diminishing. Negotiations on behalf of the entire European research community can counter this trend through "economies of scale", leading to a better negotiation position for research-relevant concessions, e.g. data ownership, reduced GDPR risk, or free data egress;
- It ensures good technical integration with the EOSC platform, making it easier to use EOSC data with commercial services and EOSC services with commercial data;
- It enables a continuous strategic dialogue with suppliers at a high level, influencing the direction in which services and business models develop;
- Conducting one procurement for the entire EOSC user community rather than thousands (institutions) or dozens (countries) of separate procurements;
- It improves access to research-relevant commercial services for all researchers in the EOSC-community[34].

As with the portfolio of centrally financed services, there will need to be a strict portfolio management process e.g. by mandating an SAC, to decide which services should be jointly procured on behalf of the whole European R&E community based on pre-defined criteria and principles.

The main cost for establishing and maintaining this portfolio of agreements is the organisation and administration of regular joint procurement exercises to establish the (framework) agreements and the subsequent effort required to manage these agreements. This includes for example demand assessment, demand aggregation, executing a large joint public procurement through a central procurement body, contract management and contract

[32] This experiment was conducted by the GN4 project.

[33] EOSC-hub deliverable D12.3 identifies and evaluates business models for such procurement and discusses opportunities and issues to address. Available from https://www.eosc-hub.eu/deliverable/d123-business-models-and-procurement-evaluation-and-recommendations.

[34] Experience from the OCRE adoption funding projects points to a likely disconnect between how researchers value the use of commercial cloud services and how e.g. central IT departments view the necessity of providing this access

EOSC Association AISBL
Rue du Luxembourg 3, BE-1000 Brussels, Belgium
+32 2 537 73 18 | info@eosc.eu | www.eosc.eu
Reg. number: 0755 723 931 | VAT number: BE0755 723 931

deployment to the thousands of organisations in EOSC. The cost per class of service is estimated to be in the order of magnitude of € 1 million/year[35].

These agreements only *enable* service consumption; financing the service usage itself follows the mechanisms of the portfolio of centrally financed services (if a service is part of that) or the mechanisms detailed in the brokerage of services.

A key challenge for the portfolio of agreements with commercial services is ensuring their availability for all EOSC participants. Joint European procurement is anchored in the EU procurement directive which unfortunately is not supported by all countries associated to Horizon Europe. This means that there are a number of countries outside the EU but firmly inside the EOSC target group which will not be able to use the agreements without additional measures being taken. There are ways to work around this which need to be further explored. This may be considered in the Task Force's future work.

4.3 Brokered not-for-profit services

This category constitutes the true marketplace of EOSC, where thousands of public sector EOSC participants can offer services to each other. The marketplace facilitates service findability, service transactions and cost recovery mechanisms which scale to a myriad of transactions while keeping the services free at the point of use for researchers. This category includes both horizontal and thematic services.

4.3.1 Possible remuneration mechanisms for the brokered services

This raises the question of what mechanisms should be put in place to facilitate the exchange between service providers and users, and the proper reimbursement to the provider of the costs incurred. The mechanism should encourage service providers to join the marketplace and provide the best-quality service possible at the best price, while allowing researchers to seamlessly access the service they need.

Previous projects[36] have already started to explore possible options such as a **voucher/token model**. Other options could include **direct payment, a subscription or freemium model** and others. A more in-depth analysis is required, to explore which mechanisms are really feasible and whether different mechanisms could potentially coexist. It is also important to note that some of these mechanisms require the respective research institutions to act as an intermediary (see Figure 2 below) whereas others, e.g. the voucher/token model, would allow a more direct transaction.

Furthermore, it will need to be explored whether a central EOSC entity will always need to act as broker for all financial transactions or whether the money can flow directly from the institution to the service provider.

[35] Calculation based on experience from OCRE project.
[36] For example OCRE and GÉANT projects.

EOSC Association AISBL
Rue du Luxembourg 3, BE-1000 Brussels, Belgium
+32 2 537 73 18 | info@eosc.eu | www.eosc.eu
Reg. number: 0755 723 931 | VAT number: BE0755 723 931

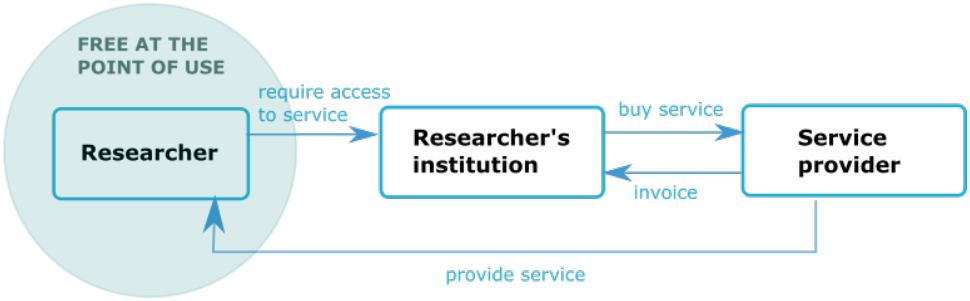


Figure 2: Possible financial transaction mechanism, e.g. via subscription model

Two major challenges (VAT and procurement directive) identified as potentially posing serious difficulties to the establishment of an ideal Exchange are discussed in the next section[37].

4.4 EOSC Exchange challenges

The organic growth[38] of European digital research infrastructures and services (e-Infrastructures) in the last 30 years and all its peculiarities influences the future development potential of EOSC. To realise a sustainable marketplace for EOSC Exchange, it is vital to better understand the landscape's point of departure, the key actors and their motivations, and the current limited funding structures within which they operate.

In recent years member states (MS) and the European Union (EU) have invested substantial resources in e-Infrastructures innovation, but sustained growth and expectations around EOSC and Open Science call for new approaches. The environment is complex. e-infrastructures can be divided broadly into two archetypes:

01. Generic e-Infrastructures (e.g. storage, compute, internet connectivity, AAI, etc)
02. Community-specific e-Infrastructures (e.g. instruments, databases, repositories, HPC architectures and usages, certification services, as well as high-level subject-specific support).

These two e-Infrastructure archetypes are often pursued by different actors and are frequently funded differently. The degree of European cooperation within the two is very distinct in nature, extent, and quality. Such e-Infrastructure actors can include:

A. e-Infrastructure provider organisations (at local, national or European levels)
B. Well-organised thematic research communities (within universities, regions, through national entities or European organisations, e.g. ESFRI, ERIC, Centres of Excellence, EIRO, …)

In addition, there are also strong collaborative networks of Open Science-driven data professionals, including researchers, which form a valuable part of the landscape whether

[37] Details on how procurement and VAT work can be found in Appendix B.
[38] Piecemeal or bit by bit growth, rather than growth according to a single overall plan.

EOSC Association AISBL
Rue du Luxembourg 3, BE-1000 Brussels, Belgium
+32 2 537 73 18 | info@eosc.eu | www.eosc.eu
Reg. number: 0755 723 931 | VAT number: BE0755 723 931

through EOSC or social infrastructures such as the Research Data Alliance[39] internationally, or through data stewards on national and local levels. These are not the focus of the comments in this section however.

Funding for European e-Infrastructures can stem from institutional, regional, national, European or international sources. Whereas some more generic infrastructures and services are funded through national funds, others compete for research grants to set themselves up or later to stay afloat. Some infrastructures or services depend on strategic research agency funding whilst others need to look to the research community or libraries to fund work in-kind or collectively to cover their operational costs using, for example, membership models. Some e-Infrastructures need to creatively utilise a range of revenue models simultaneously to survive since there are very few stable funding mechanisms available and few funders committed to funding open infrastructures[40].

Although EC funding has a powerful influence, the majority of e-Infrastructures' funding is provided at the national level or below. Member States provide long-term funding for a range of thematic research infrastructures such as ERICs or EIROs and their associated e-Infrastructures whereas the EC so far generally focuses on grant funding projects to build infrastructures or services. Project funding is temporary which poses a sustainability problem when it is used to develop critical quality innovative services or infrastructure. This problem needs to be highlighted and addressed. MS and Research Funding Organisations (RFOs) are also often at a fair distance to EU projects before, during, and after their completion. The stability of e-infrastructures can therefore be uncertain on many levels from the outset, with thematic services such as cluster projects a case in point. In short, EC funding, in principle, covers development but not operational or maintenance costs and thus leaves a big gap in the system that is yet to be filled. The EC sometimes makes efforts to sustain flagship projects by funding follow-up projects, but this is generally not the rule, and only applies to a happy few. Member States could also do more to support e-infrastructures, whether new, in development, or mature.

Current conditions for funding can inadvertently stand in the way of funding e-Infrastructures. For the EOSC Exchange to work and be sustained, funds need to be found to meet the marginal costs of the additional cross-border consumption of services arising through EOSC[41]. However, while e-Infrastructure provider organisations can be willing to allocate staff for cross-border activities if these are funded by the EC or other funding bodies, these organisations and some funders are unable to, or are not used to, invest in, co-finance or transfer funds to facilitate cross-border generic or subject-specific service provision.

[39] See https://www.rd-alliance.org/about-rda for more information.

[40] One initiative which is trying to change this is SCOSS, see https://SCOSS.org.

[41] This and related observations about the shortcomings of the current funding of the e-Infrastructures landscape are discussed in EOSC-hub *Briefing Paper on Cross-Border Services* https://www.eosc-hub.eu/publications/briefing-paper-provision-cross-border-services and in deliverable D2.5 *Final Governance and Sustainbility Implementation Roadmap* https://www.eosc-hub.eu/deliverable/d25-final-governance-and-sustainability-implementation-roadmap

EOSC Association AISBL
Rue du Luxembourg 3, BE-1000 Brussels, Belgium
+32 2 537 73 18 | info@eosc.eu | www.eosc.eu
Reg. number: 0755 723 931 | VAT number: BE0755 723 931

Furthermore, whilst member states financing European research infrastructures are willing to allocate both staff and national funds to operate joint European and cross-border e-Infrastructures which support thematic research communities, they can be reluctant to co-finance generic European e-Infrastructures. There is a clear mismatch between MS and EC commitment and funding when it comes to the question of why, how and what to fund. There are historical, cultural, legal and political reasons for these funding challenges, which need to be better understood and considered when finding an optimal sustainable programme for EOSC in the future.

This complex funding landscape is far from ideal for the long-term sustainability of EOSC. The current model in which national or EU funding is generally restricted to early-to-middle-stage innovation needs to change to keep pace with the rapid changes in the landscape. The question of who is responsible for maintaining adolescent or mature services that serve the public good has to be clarified, since currently it seems to fall outside the remit of public funding organisations. The EC and MS can be more efficient and less wasteful in what they fund in future by making concerted decisions on funding the operation of excellent services and infrastructures. To allow these to be identified, it is essential to involve the research communities as their ultimate users. This could be facilitated by implementing a Scientific Advisory Board as discussed in section 4.1.1, reporting to the funders and reviewing the quality of key services, their value for and overall use by the community/ies. Funding based on peer-reviewed input from the SAC would allow research needs to be better met by targeting funding to maintaining (sustaining) investments in innovation. For EOSC to be sustained, the EC, Member States, Research Funding Organisations and Research Performing Organisations need to review their existing funding mechanisms and explore and experiment with new funding mechanisms to help sustain a growing and maturing research ecosystem for science and society.

From the above it becomes evident that EOSC has challenges in aligning MS and EU strategic, operational, and financial commitments. It's about co-funding the same agreed activities, as opposed to funding separate activities, differently. The EOSC Partnership aims to implement the collective Strategic Research and Innovation Agenda (SRIA). The EOSC Partnership is said to be co-funded by the MS and EU partners, i.e. through EU (Horizon Europe) and MS contributions - via the EOSC Association member organisations - currently termed the Additional Activities Plan (AAP). The TF observes that AAP national contributions at this time typically do not reflect genuine co-funding towards the implementation of a joint EOSC strategy but rather consist of already allocated national funding (often targeting the transition to Open Science), which is rebranded as supportive to EOSC deployment. The potential impact a continuation of this dynamic has needs to be further investigated by the TF. The AAP should be actual additional activity, implementing national and institutional level deployment investment to make EOSC a success.

EOSC Association AISBL
Rue du Luxembourg 3, BE-1000 Brussels, Belgium
+32 2 537 73 18 | info@eosc.eu | www.eosc.eu
Reg. number: 0755 723 931 | VAT number: BE0755 723 931

4.5 EOSC Exchange boundary conditions

Certain boundary conditions need to be met if the Exchange is to grow into a smoothly functioning, dynamic single, borderless marketplace for research services and tools, providing public sector service consumers access to services offered by both public and private providers.

In particular the following issues will need to be addressed:

- **Mandate:** currently, national service providers and institutions don't typically have a mandate to provide or purchase community-services outside well-established boundaries, especially against payment, and indeed the majority of publicly funded providers in the current EOSC community face legal and governance restrictions to their ability to provide their services on the entire European public-sector research market. Where cross-border service delivery happens it's typically agreed among and financed by subject-specific communities. Only very rarely are cross-border services of a generic nature delivered, and if it happens it is established top-down from well-anchored national strategy and policy for a specific purpose, e.g. RIs, EuroHPC JU.
- **Political funding choices:** even if organisations do have a mandate to provide services outside their primary geographical area of action, there are cultural and political incentives effectively blocking this from happening: national and institutional funding is to be used only on national and institutional service production. There are exceptions to this rule where the benefits to be gained are very large, or where it is impossible to achieve a result alone (again EuroHPC JU, GÉANT, or scholarly communications initiatives such as Datacite[42]), but those remain exceptions. A rich and dynamic single European digital research service market will not be achieved unless it is a nationally supported policy, with goals allowing national/institutional funding to be used for services outside the national/institutional domain.
- **Culture change:** the current model of national service provisioning for research has been built-up over 30 years, and changing it will require strong national top-down commitment as well as time, changing protectionist incentives and a continuous systematic approach if it is to happen.
- **Institutional survivalism and national protectionism:** the current model of national service provisioning can present a threat to staff, and for good reason, by superior, cheaper, better services from other European service providers. Also however the return on national investment in infrastructures is less predictable if researchers are able to shop where they want, negatively impacting member states' ability to steer infrastructure development and, more broadly, research and research policy. Protectionism in the e-Infrastructures market must be addressed.
- **Procurement:** public sector entities should ideally be able to use services offered through the Exchange without having to conduct a public procurement process. As

[42] See https://datacite.org.

EOSC Association AISBL
Rue du Luxembourg 3, BE-1000 Brussels, Belgium
+32 2 537 73 18 | info@eosc.eu | www.eosc.eu
Reg. number: 0755 723 931 | VAT number: BE0755 723 931

these entities typically are subject to public procurement rules, a public institution in country A cannot simply buy a service from a public institution in country B unless a long-term collaboration for provisioning of services is established. In the absence of this, cross-border service provisioning against payment between organisations will be hindered in practice. In an ideal situation, the consumption of research services by public sector entities would be exempted from public procurement rules if these services are brokered through the EOSC Exchange. The details on how the EOSC Exchange can act as a procurement-free zone for public-sector EOSC participants should be the topic of a study carried out by the EOSC-Focus project.

- **VAT:** cross-border VAT is a complex subject, with VAT due in the country of service consumption and service provisioning by public sector entities VAT-exempt in some cases/countries, while non-exempt in others. In an ideal situation, VAT would not be an issue for services brokered through the EOSC Exchange. In absence of VAT-exemption, clear guidance for providers on how to deal with cross-border VAT, and practical facilitation of VAT-payments, should be part of what the Exchange does for providers.
- **Transaction support:** should the marketplace provide full logistical support for transactions between all potential actors in the marketplace, from contracting to invoicing for brokered community and commercial services and offering support for applying to centrally funded services? Should a central actor be involved in invoicing all transactions? Should B2C transactions be supported?
- **Ownership:** An EOSC marketplace and catalogue are currently being developed in the EOSC Future project, accessible through the EOSC Portal. Their future ownership is closely linked to decisions on the future EOSC legal entity, as is the case for other components of EOSC. The decisions made on the legal form of the EOSC may have consequences for transactions which take place in EOSC and for invoicing, VAT calculation and other dues and taxes.

EOSC Association AISBL
Rue du Luxembourg 3, BE-1000 Brussels, Belgium
+32 2 537 73 18 | info@eosc.eu | www.eosc.eu
Reg. number: 0755 723 931 | VAT number: BE0755 723 931

5. Towards financial scenarios for the EOSC Data Federation

5.1 Current situation

Until now, the adoption of FAIR practices and the federation of e-infrastructures to share research data in Europe has mostly taken place following community (discipline-specific) needs. The *FAIR Lady report* indicates that research data infrastructures, the underlying physical network, and the costs of making data FAIR and EOSC-compliant, need to be appropriately funded, noting that the latter are frequently not included in research budgets[43]. This is at odds with discipline-specific data federations, which have usually found ways to ensure their long-term sustainability and avoid loss of data by successful mobilisation of funding from various sources (including MS, regional, institutional and individual membership fees, or in-kind contributions). The funding model chosen depends on the organisational details of the research community (top-down *vs* bottom-up, or something in-between) and on its maturity (short *vs* long history), among other factors, leading to different paths towards sustainability.

5.2. Data Federation use cases: general considerations

Four use cases[44] providing examples of data federations, proposed by members of the Task Force, were considered:

- Blue-Cloud: thematic marine cloud developed as part of the H2020 'The Future of Seas and Oceans Flagship Initiative' for the research, monitoring, evolution and management of the marine environment, and for assessing various activities of the economy related to it (e.g. fish stocks and biodiversity, disaster management, tourist industry…)
- CESSDA ERIC: this European Research Infrastructure Consortium works to provide access to social science data and metadata for the benefit of both science and society
- Covid-19 Data Platform: joint effort by the EC and the European Bioinformatics Institute at EMBL (EMBL-EBI), together with MS and other research partners created to share Covid-19 related data and findings
- DiSSCo: currently on its way to become an ERIC, DiSSCo intends to overcome the current fragmentation of Natural Sciences Collections (NSC), i.e. scientifically classified catalogues of species, organisms, rocks, etc., building on existing databases, services and standards, by federating them.

Whilst these use cases don't provide a comprehensive analysis of the landscape, they nonetheless provided valuable examples of different ways in which data federations are created and funded. Existing experience with the development of data federations, including consideration of these use cases, allows several general observations to be made:

[43] See the *FAIR Lady Report*, page 7.
[44] See Appendix C for details.

EOSC Association AISBL
Rue du Luxembourg 3, BE-1000 Brussels, Belgium
+32 2 537 73 18 | info@eosc.eu | www.eosc.eu
Reg. number: 0755 723 931 | VAT number: BE0755 723 931

- **Data and services must go together in the design of EOSC, despite the fact that their financial scenarios are different.** Most users expect that data is retrievable together with services and applications. Although there are researchers who develop applications based on (or that make use of) available data, long-term funding is usually expected to generate both data and the related services (i.e. how to use the data).
- **Federating data implies interoperability between five levels of aggregation:** Existing data in repositories needs to be linked with EOSC and to be discoverable. However, this is not a trivial question, since a given set of data can be included in EOSC via different workflows, depending on the ownership, original provider, aggregator, or funding body. Five possible levels of aggregation can be identified: institutional, national/regional, European, thematic and international. This can generate duplication of efforts and discussions on who is providing the data and services to EOSC.
- **Data federation at EOSC level:** EOSC must consider existing data federations and repositories on different levels that comply with the FAIR principles and make them discoverable on its portal. For the time being, the Task Force assumes that EOSC will rely on existing repositories and federations[45].
- **EOSC within the twelve Common European Data Spaces or other initiatives (e.g. Gaia-X or the Global Open Science Cloud[46]):** The EC aims to have 12 Common European Data Spaces[47], most of them of thematic nature, EOSC among them. Although they all share some characteristics, EOSC's multidisciplinary scope is a bit different. Unlike others, it includes actors from other data spaces among its stakeholders. For sustainability and efficiency it will be important to clarify how EOSC fits and positions itself among the data spaces and what the synergies will be. One possible opportunity is European Smart Middleware for Data Spaces -Simpl[48], for which a proof of concept may be available by the time of publication of the Task Force's final report in 2023.
- **Positioning of EOSC in the international sphere as an incentive for more funding. European contribution, competitiveness and visibility on the international scene.** There exist data federations beyond Europe with a global scope, most of them with participation of European organisations. The presence of data from Europe in those federations following international data standards improves visibility and competitiveness in the international scene, and expands the funding opportunities. EOSC should therefore establish itself in the international scene, and address any

[45] The Task Force Long-term Data Preservation is currently preparing a document which will inform this Task Force on this matter.
[46] The Global Open Science Cloud or GOSC is an initiative led by CODATA. For a description of its background and aims, see https://codata.org/initiatives/decadal-programme2/global-open-science-cloud/
[47] See https://digital-strategy.ec.europa.eu/en/library/staff-working-document-data-spaces
[48] See https://digital-strategy.ec.europa.eu/en/news/simpl-cloud-edge-federations-and-data-spaces-made-simple

EOSC Association AISBL
Rue du Luxembourg 3, BE-1000 Brussels, Belgium
+32 2 537 73 18 | info@eosc.eu | www.eosc.eu
Reg. number: 0755 723 931 | VAT number: BE0755 723 931

barriers in the way of cross-border collaboration, including access to international funding mechanisms.

- **A full assessment of the financial needs of establishing and running the EOSC DF requires consideration of costs other than the creation of the "overlay" or catalogue:** costs associated with the establishment of data management infrastructures include staff; long-term storage and maintenance of research outputs; data stewardship and other categories. To ensure long-term sustainability, it is important that MS/AC consider these as eligible within national funding schemes, which will require policy changes at MS/AC and EU level.
- **EOSC DF must be created using existing infrastructures and thematic ecosystems without duplicating efforts.** Discussions with infrastructure managers and funders still need to take place about the costs and required commitment from different stakeholders that each of the possible solutions involve.

5.3 Possible Data Federation architecture models

In our initial analysis several possible options were identified for the EOSC DF:

1) **"Overlay":** to enable data discovery from any provider, EOSC could put in place an intermediary (software) layer, centrally managed and maintained by EOSC, that ensures data interoperability to the highest degree possible. This would come "on top of" (i.e. in addition to) any DF layers which already exist, e.g. in a specific scientific domain or geographical area.

2) **Metadata catalogue:** Instead of linking the data, a system to track *metadata* (i.e. not data itself, but their attributes or characteristics), organised into a catalogue, could be adopted by EOSC. This catalogue could then be used to achieve data interoperability, since having a coherent and consistent metadata classification scheme for data of very different nature allows to find and eventually use diverse types of data[49].

3) **Catalogue of data providers and platforms:** If a common interface to access the (meta)data cannot be implemented, an alternative solution would be to compile a list of all data providers and the platforms they use, indicating how to access single or combined data sources depending on the scientific discipline or geographic area.

4) **Natural evolution**: In the absence of the EOSC, existing data infrastructures will remain as the go-to sources for researchers. Convergence between them would continue to occur but it would be uncoordinated and slow, involving significant costs and implying the extension of the current suboptimal (inefficient) use of resources and siloed data landscape.

[49] OpenAIRE's Research Graph (https://graph.openaire.eu/) is a possible candidate which could provide *part of* the basis for a centralised EOSC metadata catalogue, with FAIRsharing providing its (manually) "curated database, policy and standard metadata" as "an authoritative source".

EOSC Association AISBL
Rue du Luxembourg 3, BE-1000 Brussels, Belgium
+32 2 537 73 18 | info@eosc.eu | www.eosc.eu
Reg. number: 0755 723 931 | VAT number: BE0755 723 931

Due to time constraints, the Task Force has so far only been able to discuss and analyse the "overlay model" as a financial scenario in greater detail. The other models described above may be considered in more detail in the next phase of the Task Force's work.

5.4. Conclusions from the use case comparison - additional costs for the EOSC Data Federation

One feature common to all four use cases is that they have established some kind of "overlay", either in the form of an international organisation responsible for providing the necessary infrastructure for community-specific, federated data (CESSDA and DiSSCo), or in the form of a platform which enables the collection, discovery, retrieval and sharing of research data (Covid-19 Data Platform and Blue-Cloud). The additional costs involved in creating an overlay are discussed in this section.

In all four use cases, the costs of data resources (at national, European or community level), and operation and maintenance costs, are mainly covered by various forms of member state (MS) funding (membership fees from MS to international organisations, funding from public and private foundations, etc). EU and national project funding is mostly used to cover costs related to creating connections between data resources and facilitating accessibility.

This suggests that the additional costs for the EOSC DF should be split into two categories: costs related to the establishment and maintenance of the overlay (which would be the visible part of the EOSC DF), and costs at European thematic, national/regional and institutional levels. Since all these other "underlying" levels will eventually feed their data directly or indirectly into EOSC, the overlay of the EOSC DF can only be successful (and sustainable) if the parts that constitute it also have a solid and financially sustainable basis.

5.4.1. Additional costs of the EOSC overlay

If an EOSC DF overlay were to be established, whether as an international organisation or a platform, the use cases indicate that establishing and operating it will entail several additional costs associated with the discovery, access and retrieval of either metadata or actual data sets and products using a common interface. The final decision will depend on the desired complexity (from a simple metadata catalogue to a portal that allows active submission of datasets or other functions; tools for mapping and viewing the location of datasets (as included for example in Blue-Cloud DD&AS, see Appendix C), or a shopping mechanism to facilitate users to compose shopping baskets, to requests for datasets from multiple data infrastructures) and of what is actually viable from the constituents of the federation.

Once in place, the EOSC overlay would also generate maintenance and operational costs, mainly but not exclusively related to management of IT resources. Taking the Covid-19 Data Platform as an example, sustaining the platform requires about half the level of resource on an ongoing basis as the set-up phase).

EOSC Association AISBL
Rue du Luxembourg 3, BE-1000 Brussels, Belgium
+32 2 537 73 18 | info@eosc.eu | www.eosc.eu
Reg. number: 0755 723 931 | VAT number: BE0755 723 931

5.4.2 Costs for aggregation levels below EOSC: EU, thematic, national, and institutional dimension

As indicated at the beginning of this chapter, linking data to EOSC comes with a significant extra effort and cost for data providers. There are already many thematic RIs in operation with sustainable MS funding and a clear mandate to provide community-specific FAIR and open research data, such as EMBL and other EIROs, CESSDA and the other ERICs, or DiSSCo. However, this mandate and the financial resources provided are not sufficient to cover the costs of creating links and establishing interoperability at EOSC level.

Data-sharing practices have been used successfully for more than eighty years in areas such as astronomy, bioinformatics and molecular biology, thanks to a sustained community effort driven by the general understanding that research in this area is necessarily built on the reuse of data; other disciplines have only recently started to discover the benefits of data reuse. Despite unequal levels of development, most communities share similar problems when using data for inter- or trans-disciplinary research, such as lack of clear responsibilities in the data curation process, especially in the description of the information contained in the data (i.e. creation of metadata).

To achieve interoperability between data catalogues in the EOSC DF, it is important to recognise and plan for extra costs, both relating to connecting different data resources to EOSC and to making data FAIR. These extra costs, which concern all the different levels of aggregation, have been identified and grouped in table 1.

Making data FAIR	***Making experiments reproducible***
○ Operational Costs ■ Data Stewardship ■ Control Systems (for acquiring metadata) ■ Storage (curation, hardware, software) ■ Data transfer (network, software) ■ Access to computing (identification, prioritisation, computing resources e.g. for data reduction) ■ Making data findable (assignment of DOI, catalogue, search tools) ○ Development costs ■ for capturing metadata (e.g. electronic logbooks, metadata models) ■ for improving data reduction ○ Costs related to converting legacy data into FAIR data	○ Development of workflows and software ○ Development of data analysis environments ○ Development and maintenance of software catalogues ○ Development and maintenance of computing capacity to run data analysis ○ Development and maintenance of data transfer protocols

EOSC Association AISBL
Rue du Luxembourg 3, BE-1000 Brussels, Belgium
+32 2 537 73 18 | info@eosc.eu | www.eosc.eu
Reg. number: 0755 723 931 | VAT number: BE0755 723 931

Ensuring long-term access to data[50]	***Federating data to EOSC***
○ Storage and archival costs for sustainable repositories ○ Costs to update metadata and APIs ○ Costs to connect endpoints and make data findable via data catalogues across Europe ○ Cost of all tools and services for making data FAIR	○ Creating links and enabling interoperability between data resource and EOSC (enabling feeding into EOSC) ○ Operational costs related to EOSC DF (compliance with interoperability)

Table 1: Additional Costs Involved in achieving interoperability between data catalogues in the EOSC DF

It is important to highlight that some of these aspects, such as data stewardship, are required at all levels (European, thematic, national, institutional). Others, however, require a clear assignment of responsibilities, which is yet to be done. One example of this is data preservation: the *FAIR Forever* study by the Digital Preservation Coalition[51], found that, once the data leaves the institution where it was created to be stored (e.g. by transfer to an "upper level" of aggregation, such an overarching institutional system, or a discipline-wide or national infrastructure), researchers stop being responsible for its management. If the responsibility for data preservation across the levels of aggregation is not clearly defined, the chain may be broken at some point, which may cause data to be lost. Even though institutions know about data stewardship, it 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" (*FAIR Forever*, p. 23). The study, in line with the *FAIR Lady report*, concludes that roles, responsibilities, and accountabilities for preservation in EOSC are opaque and should be clarified, together with the associated salaries and funding streams; otherwise, there is a risk to data, reputation and sustainability that will prevent EOSC achieve its long-term goals (p. 39). The Financial Sustainability Task Force is in regular contact with the Task Force on Long-term Data Preservation to address this issue[52].

5.4.3 Costs of de-duplication: One data set, but multiple entries in EOSC

Multiple entries on a single entity are an important issue when combining catalogues into one EOSC catalogue or portal. They arise when the same entry is being harvested from different sources. This cannot entirely be avoided due to the fact that the data catalogue of a data provider is harvested by search engines and several other providers who act as metadata aggregators. These, and possibly also the original source itself, are then harvested by the EOSC

[50] The Task Force on Long Term Data Preservation (see https://www.eosc.eu/advisory-groups/long-term-data-preservation) may also provide comments on this aspect in a similar timescale to the present progress report.
[51] Currie, Amy, & Kilbride, William. (2021). *FAIR Forever? Long Term Data Preservation Roles and Responsibilities, Final Report (Version7)*. https://doi.org/10.5281/zenodo.4574234.
[52] For more details on their scope and aims see https://www.eosc.eu/advisory-groups/long-term-data-preservation.

EOSC Association AISBL
Rue du Luxembourg 3, BE-1000 Brussels, Belgium
+32 2 537 73 18 | info@eosc.eu | www.eosc.eu
Reg. number: 0755 723 931 | VAT number: BE0755 723 931

catalogue resulting in multiple entries for a single entity. De-duplication techniques are implemented to solve this issue.

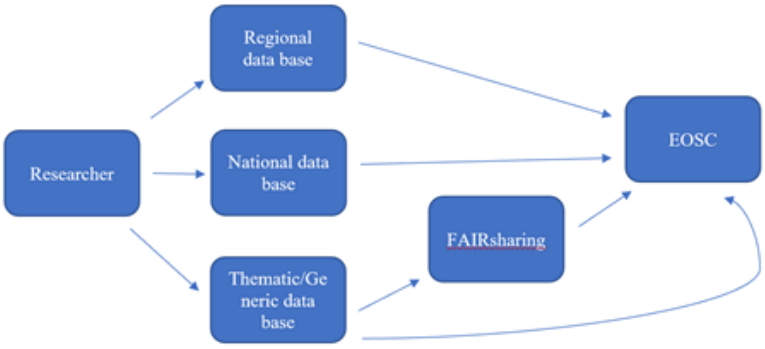


Figure 3: De-duplication need

Another important aspect which should be dealt with is that to achieve higher visibility, many funding bodies want to be the official way to EOSC - some funders even make it a condition for applicants. This leads to a duplication of efforts: researchers need to make the same data available in different repositories because different funding bodies insist on controlling the access route for the same data set/RIs. As a side note, RI catalogues can also benefit from pulling entries from the EOSC catalogue to enhance their services and show the available resources to their users. Standardised interoperability protocols exist to accomplish such push and pull mechanisms (e.g. OAI-PMH[53])[54]. Costs arise from implementing these interface standards into the platforms.

While de-duplication makes it technically possible to interlink the content through PIDs and other FAIR data tools to reduce duplication, this does not solve the problem because the political issue may persist: funders might insist being the main aggregator, rather than relying on the researchers to decide the most suitable data repository and accepting that the funding body is part of the metadata schemes. A more cost-efficient approach for displaying relevant information on funders' websites would be the use of widgets that pull and display the relevant information, potentially saving additional efforts to maintain and produce databases that duplicate information. An additional effort (with related additional costs) could therefore be assumed by the EOSC overlay to curate duplicated entries which cannot technically be avoided with de-duplication tools, while ensuring that the visibility of funders is not compromised.

5.4.4 Costs of harmonisation: EOSC, thematic and other data portals

Interlinking community data to EOSC is often perceived as an additional layer where users need to be convinced of the need to make the extra effort of linking their data in an EOSC-compliant way, especially where community data sharing tools already exist. This raises

[53] Open Archives Initiative Protocol for Metadata Harvesting https://www.openarchives.org/pmh/
[54] The EOSC Association Technical Challenges Advisory Group task forces will also be addressing such issues.

EOSC Association AISBL
Rue du Luxembourg 3, BE-1000 Brussels, Belgium
+32 2 537 73 18 | info@eosc.eu | www.eosc.eu
Reg. number: 0755 723 931 | VAT number: BE0755 723 931

important questions about the future: once a model for federating data into EOSC exists, what will happen to the current thematic data portals (i.e. repositories)? Should such general and thematic aggregators and EOSC exist in a complementary way without being perceived as overlapping or duplicating efforts? In the current landscape, generic and thematic data portals are *the* places to store data and metadata, because they follow discipline-specific rules and ensure they are implemented and followed by others, and, most importantly, they provide and regulate access to the data. At any rate, an important cost driver for data aggregation of any future interlinked repository network is harmonisation (i.e. alignment) of metadata schemas (catalogue of datasets and services) and the associated access procedures. Appropriate funding is needed also for certification and validation of repositories, and to cover the costs of operation and maintenance.

It is important to understand that the landscape of data aggregators which may contribute to the EOSC Data Federation is itself in a big transitional process which must acknowledge the historical legacy of those data aggregators. Together with the developments of RIs and elsewhere, it is unlikely that either thematic data portals and services or more generic aggregators, such as the various ERICs, Blue-Cloud, or the Covid-19 Data Platform, will disappear and be replaced by EOSC in the near future. Hence, it will come down to interoperability and collaboration between them and EOSC, and further investigation into whether a closer integration into a possible EOSC overlay would lead to financial advantages will be needed in the future. This sort of preparation for (greater) federation is a non-trivial exercise which requires a clear vision for the long-term architectural set up of the EOSC DF.

5.4.5 Costs of legal and ethical aspects, incl. sensitive data

Discussion with the Task Force Long-term Data Preservation has shown that legal and ethical recommendations and requirements also drive costs. It is especially challenging for smaller institutions—which often lack the necessary resources—to conform to the additional layer of EU legislation and best practices, resulting in them being left out of data federations and associated services. While EOSC obviously advocates Open Science, there are many legal barriers and ethical issues to deal with, such as those posed by the General Data Protection Regulation (GDPR) and by intellectual property rights and copyright, among other legislation. The cost this entails must be considered for the overall sustainability of EOSC.

The cross-border mapping of different countries, EU or international legislation for such a large multidisciplinary network as EOSC with multiple and sometimes unclear responsibilities is far from trivial. Additional costs can appear in the future from IP infringement claims towards associated portals and services. Many of these claims are not even legitimate but those accused do not have the means to prove it and end up paying non-negligible fines for take-down orders or legal cost for putting up a defence.

5.4.5.1 Federating sensitive data

A particular challenge is the federation of specially protected or sensitive data. A good example for the successful implementation of a “sensitive data federation” is the European

EOSC Association AISBL
Rue du Luxembourg 3, BE-1000 Brussels, Belgium
+32 2 537 73 18 | info@eosc.eu | www.eosc.eu
Reg. number: 0755 723 931 | VAT number: BE0755 723 931

Genome-phenome Archive (EGA)[55], which is one of the main pillars of the Covid-19 Data platform. The EGA is a resource for permanent secure archiving and sharing of all types of potentially identifiable genetic, phenotypic, and clinical data resulting from biomedical research projects. This data is subject to participants' consent agreements, so sharing is restricted to researchers for specific research purposes. Operating a distributed data access model, the EGA provides the infrastructure and services for secure data archiving and distribution for over 1,500 different Data Access Committees (DACs).

EGA services include data submission, discovery, and access to the global research community. The provision of these services for such sensitive data comes with additional efforts and thus additional costs to meet the ethical and legal standards. This includes, of course, the necessary legal expertise, but also the development of specific technical solutions with regards to API, controlled vocabularies and accession as well as secure, efficient and robust download and decryption of data in order to ensure the security of the data as well as the users' need for interoperability. Moreover, the EGA follows a strict security strategy including regular risk assessment, risk mitigation, identity and authorisation management, audit logs as well as cryptography, communication security, and data integrity.

5.5 How to address these challenges

It has been suggested[56] to have expertise or content on legal and ethical data within EOSC. Such a solution could be envisaged as a help desk or central contact point. Legal questions arise when sensitive data is transferred from one country to another, spanning two different legal frameworks. This would represent a large data mining activity, with high costs for harmonisation and developing crosswalk-schemas of the legal frameworks. Thematic policies platforms are being built (e.g. in the biodiversity domain), but remain largely at a conceptual phase.

Funding for conformity and dealing with legal and ethical aspects should probably come from local national and regional authorities towards their EOSC data providers and be part of the overall governmental support, including the alignment with national, EU and international regulations. A decision on the responsibilities and who pays in case of infringement claims has to be clearly defined and communicated.

The Task Force recommends that EOSC-A has a (federated) group of experts specifically dealing with legal and ethical aspects relevant to EOSC, including but not only those relating to data. In addition a contingency fund could complement an EOSC risk management framework.

In wrapping up the discussion and cost issues of the Data Federation chapter, it is obvious that the existing huge diversity of funding models brings a considerable fragmentation and

[55] https://ega-archive.org/

[56] For example, see EOSCpilot deliverable D3.6 "Final Policy Recommendations", in particular implementing action 1.4 in section 2.1, and Annex F, available at https://eoscpilot.eu/sites/default/files/eoscpilot-d3.6-v2.7_0.pdf

EOSC Association AISBL
Rue du Luxembourg 3, BE-1000 Brussels, Belgium
+32 2 537 73 18 | info@eosc.eu | www.eosc.eu
Reg. number: 0755 723 931 | VAT number: BE0755 723 931

complexity to the European RIs landscape and leads to **misunderstandings**. It is therefore advised to continue the efforts for a thorough mapping of the **curation, the quantity and the quality of the data** in RIs, together with monitoring at a national level. As a prerequisite an exhaustive list of relevant infrastructure organisations is key[57].

Once there is a clear picture of the assets (data and services), it will provide a basis for assessing what is required for EOSC, and also where there is duplication or inefficiency, which could help to inform decisions about funding.

Existing experience can also provide valuable examples which can be drawn on when considering the financial sustainability of the EOSC Data Federation. For example, financial annual statements from ERICs and non-ERICs organisations offer essential know-how on how data is transferred and used (e.g. the ELIXIR community does this directly through their national nodes). Financial plans from distributed RIs have to provide answers to:

- How to manage the percentage of the assets that they control and that they may offer to the EOSC ecosystem
- How to manage the percentage of assets that are not under their direct control and controlled by others.

Many of these RIs have been developed in recent decades through a number of research projects (EC Framework Programmes and other funding opportunities, including structural funds), which has for the most successful projects ensured funding of the implementation and construction but has not guaranteed their sustainable operation.

Successful hybrid models exist where RIs have organisational bodies containing members of the community, to which the RIs report, while at the same time being sustainably funded by a mix of funding bodies. Other research infrastructures ensure the engagement of the communities as well as member states, research organisations, or universities.

The Task Force expects to study such examples further in its future work.

6. Consultation

6.1 Consultation

As explained in the Introduction, the Task Force has not fully developed financial scenarios for the EOSC but we would like to gather feedback on our findings so far as described in this report. We wish to address primarily the EOSC Tripartite Governance, EOSC Association mandated organisations and Task Forces, and relevant experts such as those in EOSC projects, ESFRI RIs and clusters, and European e-Infrastructures.

A consultation exercise will be launched shortly after publication of this report.

[57] The Task Force will liaise with the Task Force on FAIR Metrics and Data Quality to discuss relevant needs and activities.

EOSC Association AISBL
Rue du Luxembourg 3, BE-1000 Brussels, Belgium
+32 2 537 73 18 | info@eosc.eu | www.eosc.eu
Reg. number: 0755 723 931 | VAT number: BE0755 723 931

6.2 Task Force next steps

The consultation feedback collected will help to inform the Task Force's plan of work over the next year, but from our work to date a number of possible activities can already be identified as possible areas for the Task Force to address in its next phase of work, which are listed below. The landscape in which the Task Force is working is highly dynamic however, so the precise workplan to be followed is liable to be influenced by external developments:

6.2.1 EOSC Core

The Task Force has presented in this report a financial scenario for the Core. Many of the questions it raises relate to the future legal entity, governance structure and composition chosen for the Core. These aspects are beyond the scope of the Task Force and we do not expect, unless requested to do so, to study them further except to the extent that they are of relevance to our further consideration of the EOSC Exchange and Data Federation.

The EC is however expected to shortly launch a procurement for the EOSC Core, which the Task Force may comment on in the coming months.

6.2.2 EOSC Exchange

For the Exchange, the Task Force presents in this report the elements of the Exchange for which financial scenarios are required and presents the main challenges and boundary conditions it has identified in relation to them. Several areas have been identified which require further research and analysis, including:

- The value proposition for the centrally financed portfolio of services, possibly making use of case studies, and perhaps including some quantification of the amount of consumption the centrally financed portfolios need to support. Case studies may be used to gather evidence
- Facilitation and remuneration mechanisms for brokered services in the marketplace, including whether they will always require a central EOSC entity to broker financial transactions
- Detailed studies into critical aspects to support scenarios, particularly relating to VAT and procurement.
- Further development of financing models.

These points comprise numerous large and complex issues whose in-depth analysis is beyond the capacity of a volunteer Task Force so we expect to apply some prioritisation and selection. See also section 6.3 below.

6.2.3 EOSC Data Federation

For the Data Federation, in the ongoing absence of an agreed architecture, the work of the Task Force is likely to be influenced by feedback from the consultation exercise, but possible next steps include:

EOSC Association AISBL
Rue du Luxembourg 3, BE-1000 Brussels, Belgium
+32 2 537 73 18 | info@eosc.eu | www.eosc.eu
Reg. number: 0755 723 931 | VAT number: BE0755 723 931

- Analysis of the results of the EC-commissioned study on characterisation of the European Research Data Landscape and outputs from the Task Forces on Long-term Data Preservation and FAIR Metrics and Data Quality
- Exploration of other data federation models - for example meta-data catalogue, catalogue of data providers and platforms, natural evolution - and comparison with the overlay scenario
- Further investigation of the four case studies already begun or development of additional case studies
- Liaison with the Data Spaces Support Centre[58] for alignment with the Smart Middleware Platform

6.3 Support from EOSC Focus

The EOSC Focus project includes a task dedicated to business models, and also some consultancy budget, which can provide support to the Task Force. The following requests for support will be discussed with the EOSC Focus project:

1. Manage consultation exercise with key stakeholders to gather feedback on the proposals in this progress report, including analysis of the responses
2. Two pieces of specialist advice relating to the legal entity chosen for the Exchange:
    a. advice of a tax expert on the EU VAT consequences of the choice, and
    b. advice of a lawyer on the public procurement requirements which would apply to the Exchange owner(s) and users
3. Specialist analysis and advice on facilitation and remuneration mechanisms for brokered services in the Exchange and whether they will always require a central EOSC entity
4. Based on consultation feedback and under the direction of the Task Force, conduct studies to estimate the implementation and operation costs of one or more architecture models for the EOSC Data Federation.

6.4 Conclusions

As already observed, the landscape in which the Task Force is working is evolving rapidly and numerous strands of activity are taking place currently which will inform or influence our future work. The aim laid out in our Charter - to provide by 2023 a proposal for long-term financial sustainability of the main building blocks of EOSC (EOSC Core, EOSC Exchange and the Federation of data & Data Services) - remains challenging, with a large amount of work to do to study the necessary aspects relevant to developing validated scenarios and recommendations for financial sustainability of EOSC. We will work closely with the EOSC Focus project and draw on the outputs of other EOSC-relevant projects to make the best possible use of the resource and expertise available, but we also expect to scale our activities in proportion to the resource available within the Task Force.

[58] For more information, visit https://dssc.eu/

EOSC Association AISBL
Rue du Luxembourg 3, BE-1000 Brussels, Belgium
+32 2 537 73 18 | info@eosc.eu | www.eosc.eu
Reg. number: 0755 723 931 | VAT number: BE0755 723 931

Appendix A - EOSC Core governance requirements

1. Who should govern (decision making body) and fund the EOSC?		
Options	*Advantages*	*Disadvantages*
EC	European strategic focus. Possibility to make EOSC synergetic with other EU initiatives. Guaranteed inclusiveness for all EU countries. Priorities of action strongly influenced by the EC. Structural access to extensive funds from all EU Member States. Political power.	Necessity to find a legal form that can incorporate the EC in the governing body. This may exclude the possibility to participate in competitive funding calls issued by the EC
MS	Engagement of individual MS. Possibility for MS to contribute to strategy and priorities. Straightforward funding mechanism through agreed fees. Better coordination between national and European investments and improved involvement in EOSC at the national level. Community influence and governance by those who contribute financially. MS funding contributes to the success of EOSC as opposed to no funding or sole EC funding.	Necessity to select a legal form that can incorporate MS in the governing body need to find a fair contributing and governance mechanism that allows countries with very different critical mass and level of resources to be represented, negotiation needed to involve the largest possible number of MS. Need to have an incentive for membership over non-membership for MS to join and contribute. May "crowd out" participation of individual RIs or RPOs within each country? Lack of good practices where this international joined up infra community-funded model works (except for GEANT and a few others). Will protectionism and competition amongst MS deem this impossible?
Countries associated to Horizon Europe	These countries play a role in European research, therefore they could be considered on the same grounds as MS. They are consumers as others, but related to EC, so have preferential role/status. They could also financially contribute if MS states do.	The list of countries can change between MFFs, therefore if they are to be considered for the coverage of EOSC Core services costs it may be necessary to adjust contributions from other members as countries are removed or added to the list. Advantages of being MS *vs* third country in EOSC in this respect not clear.
Public organisations or organisations of public interest	Direct involvement of end users in governance. Long-term information of service providers. Some public orgs. have formal networks and are well organised and strategically aligned on OS like LERU, EUA, can impact change	Potential lack of European vision in the strategy and of coordination between national and European funding and initiatives; ineffective governance; stronger representation from some actors, steering the strategy of the entity with a limited vision

EOSC Association AISBL
Rue du Luxembourg 3, BE-1000 Brussels, Belgium
+32 2 537 73 18 | info@eosc.eu | www.eosc.eu
Reg. number: 0755 723 931 | VAT number: BE0755 723 931

	and adoption - more for governance than funding? Trusted partners.	
Companies	Effective communication with service providers. Innovative, good foresight, business acumen and tech know-how	Uneven access to information for private companies, which may violate public procurement principles (see disadvantages for public organisations). Undue influence by companies where profit making is more central than service; opposing academic values. Potential mid- to long-term lockin if part of the solution early on (dependence). Can generate lack of trust.

2. With which "granularity" should be decided who should be able to use EOSC Core services?		
Options	*Advantages*	*Disadvantages*
The EOSC involves all the EU by default	EOSC Core federates data and services from all EU countries. Citizens from any EU country can be authenticated and authorised to discover and use data and services federated in EOSC.	Possible uneven financial contribution from funders. They need to agree on providing access to services also to MS/institutions that may not contribute to the maintenance of the EOSC Core. Excludes those outside of the EU like the USA.
MS	EOSC Core federates data and services from all EU countries, but only Citizens from some EU countries can be authenticated and authorised to use the data and services federated in EOSC.	Stronger motivation for MS to contribute financially to the maintenance of EOSC Core. The advantages of OS for publicly funded research outputs not accessible or exploitable by all but only to privileged MS. DEI in Europe not addressed adequately
Public or private organisations	EOSC Core federates data and services from the organisations that are authenticated and authorised to discover and use data and services federated in EOSC.	Difficult governance; incomplete coverage at national and European level; less added value for EOSC. Previous experience (GÉANT) discourages to adopt this model

3. Who should contribute to maintaining the EOSC Core?		
Options	*Advantages*	*Disadvantages*
EC	A financial contribution from the EC would ensure its weight in the governance and make it easier for	Necessity to find a legal form that can incorporate the EC amongst the

EOSC Association AISBL
Rue du Luxembourg 3, BE-1000 Brussels, Belgium
+32 2 537 73 18 | info@eosc.eu | www.eosc.eu
Reg. number: 0755 723 931 | VAT number: BE0755 723 931

	MS with less resources and critical mass to be represented both financially and in the governance	funding bodies
MS	MS will benefit from EOSC, so they are a clear candidate to fund EOSC Core and EOSC in general. A careful agreement on the fee contributions should allow all MS to participate	Necessity to select a legal form that can incorporate the MS as funders, need to find a fair contributing mechanism that allows MS with very different level of resources to be represented, negotiation needed to involve the largest possible number of MS
Countries associated to Horizon Europe	Allowing third countries to contribute can enlarge the base of datasets available, critical for some applications; increased sustainability	Careful consideration of the inclusion of the contributors in the governance and agreements based on reciprocity or mutual interest will be needed.
Public organisations or organisations of public interest	Costs could be divided over a larger number of contributors	Very heterogeneous funding situations amongst public organisations. Impossibility to define a sustainable model. Complex national and regional mechanisms will be needed to transfer funds to organisations contributing to EOSC.
Companies	The cost could be divided over a larger number of contributors	Careful consideration of the inclusion of the contributors in the governance will be needed. Lack of added value for commercial providers to sustain the EOSC. Previous experience (GÉANT), discourage to adopt this model

4. Should the EOSC Core provide commercial services? Or should the legal entity be suitable also for the Exchange?		
Options	*Advantages*	*Disadvantages*
Yes	Source of revenue to contribute to the sustainability of the core; increased added value	The legal form chosen needs to allow performing commercial activities. Excludes participation already at the Core. Unadvisable
No	All legal forms could be applicable, no restrictions	Missed opportunities; need to find a different legal entity to govern the Exchange (increased complexity)

EOSC Association AISBL
Rue du Luxembourg 3, BE-1000 Brussels, Belgium
+32 2 537 73 18 | info@eosc.eu | www.eosc.eu
Reg. number: 0755 723 931 | VAT number: BE0755 723 931

5. How long should the long-term funding be?		
Options	*Advantages*	*Disadvantages*
Indefinite	The lifetime of EOSC is secured by regular funding	Some funders may not be able to commit for long-term. Possible legal forms to do this might be restricted. Not an advisable option, since danger exists of excluding evaluation of the value of EOSC over time.
Multi-annual	More flexible allocation of funding, possibility to synchronise with the European multi annual financial framework	Need to secure funding for operation before the end of each multi-annual period
Annual	More flexible allocation and use of funding	Lack of long term vision, high risk, high uncertainty. Nor really a long-term solution

6. What kind of benefits should the legal entity offer?		
Options	*Advantages*	*Disadvantages*
VAT exemption	Better use of the joint budget	The legal form chosen needs to foresee the vat exemption
Own procurement rules	Less bureaucracy, more effective operation	The legal form chosen needs to foresee the vat exemption

EOSC Association AISBL
Rue du Luxembourg 3, BE-1000 Brussels, Belgium
+32 2 537 73 18 | info@eosc.eu | www.eosc.eu
Reg. number: 0755 723 931 | VAT number: BE0755 723 931

Appendix B - Details on procurement and VAT

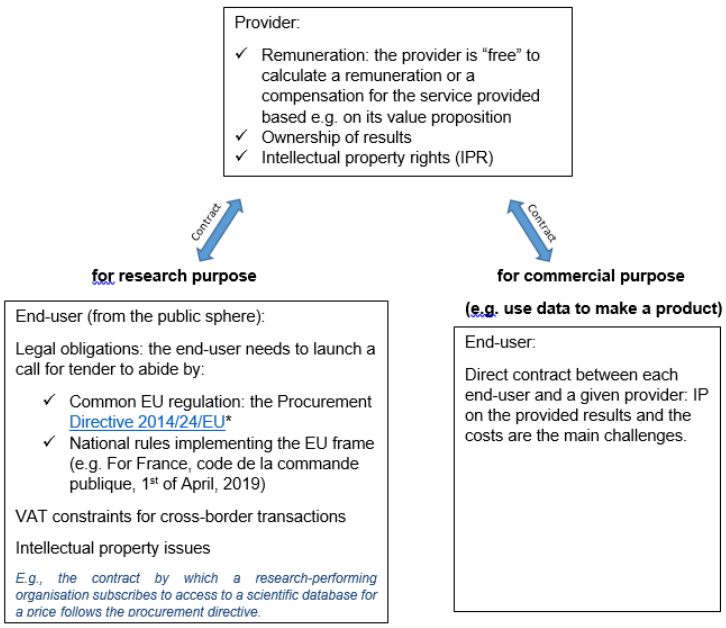
The Financial Sustainability Task Force recognises that the legal frame that applies to service providers for the establishment of a price differs from that of service buyers, which are mainly research performing organisations that have to follow public order (procurement directive: best offer/best price), whereas providers will have to pay attention to VAT/currency rates outside the euro zone if its market is in the EU.

B.1 European Procurement Directive

Compared to the funding of provision and acquisition of services through grants, procurement is a much more strictly regulated procedure, but if properly designed it guarantees that the usual conditions of best price, quality, transparency, etc. demanded by taxpayers are kept, and lowers the risk of failure and corruption. The additional regulation and necessary administration can become a hurdle for researchers and institutions. In some cases the extra paperwork needed can even lead to potential users refraining from using a given service. Procurement is a hurdle that needs to be solved. If too much administration is required to access a procurement process, researchers or institutions will refrain from using a given service, prefering to buy local services even if that solution is not as good as the one offered across the border. The consequences are mainly of two types:

1) The buyer must consider the terms of conclusion of the contract: in principle, the conclusion of a public contract must be preceded by a procedure of advertising and prior competition. This procedure allows the buyer to choose the most economically advantageous offer, i.e. the one presenting the best quality/price ratio;
2) The contract must be performed in accordance with the provisions of the public procurement code in France or in accordance with the national legislation of the MS. The execution of a public contract must respect certain obligations specific to public procurement. For instance, the modifications during the execution of a public contract are strictly framed: the national legislations list the hypotheses of authorised modifications.

EOSC Association AISBL
Rue du Luxembourg 3, BE-1000 Brussels, Belgium
+32 2 537 73 18 | info@eosc.eu | www.eosc.eu
Reg. number: 0755 723 931 | VAT number: BE0755 723 931



The European Directive 2014/24/EU[59] sets exceptions to the obligation of organising a prior advertising and competition procedure. The public procurement code[60] exhaustively lists these exceptions. European and national case law requires the buyer to have a legal argument, demonstrating that the conditions for recourse to an exception are met. Otherwise, the risk is the cancellation of the contract by the administrative judge and the buyer's criminal liability may be engaged in front of the criminal judge. If the conditions for recourse are met, the market is said to be concluded "by mutual agreement".

For illustration, article L. 2512-5 of the French public procurement code allows a buyer to conclude over-the-counter contracts relating to the provision of research and development services and for which the buyer (i) does not acquire exclusive ownership of the results, OR (ii) does not fully finance the service. The buyer will have to follow the public procurement procedure, which becomes more restrictive with increasing amount of the contract.

Public procurement law provides for three levels of procedure:

- the absence of constraints for public contracts below an amount of **40,000 € excluding tax** (this is the French threshold). However, the regulation requires the buyer to choose a relevant offer, to make good use of public funds and not to systematically contract with the same economic operator;

- the adapted procedure: intermediate procedure, it involves different terms of competition and publicity depending on the amount of the contract;

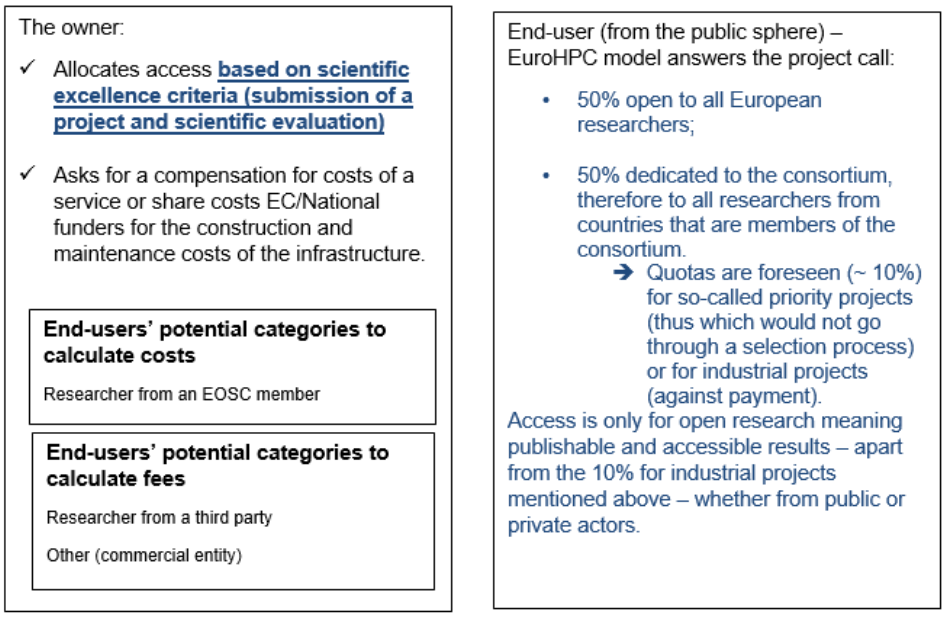
- the formalised procedure: this is the most restrictive procedure, reserved for contracts of a significant amount. It involves substantial procedural constraints.

[59] *Directive 2014/24/EU of 26 February 2014 on public procurement and repealing Directive 2004/18/EC Text with EEA relevance*, https://eur-lex.europa.eu/legal-content/en/TXT/?uri=CELEX:32014L0024
[60] https://ec.europa.eu/info/policies/public-procurement_en

EOSC Association AISBL
Rue du Luxembourg 3, BE-1000 Brussels, Belgium
+32 2 537 73 18 | info@eosc.eu | www.eosc.eu
Reg. number: 0755 723 931 | VAT number: BE0755 723 931

Depending on the estimated amount of the contract, the buyer organises the appropriate competitive bidding procedure. **Each year, the European Commission sets the thresholds above which the buyer is obliged to use one of the aforementioned procedures**.

B.2 VAT



General rules applicable in terms of VAT

- Directive 2006/112/EC modified regarding the place of services by the Directive 2008/8/EC
- National rules

To determine the applicable fiscal rules, we recommend to answer to the following questions:

- Does the transaction in question fall within the scope of the service user's VAT?
- Is the co-contractor subject to VAT in his country?
- Can the operation benefit from a VAT exemption?
- If the transaction is taxable, who is liable for VAT?
- What are the applicable invoicing rules?

B.3 How to address these challenges

The Task Force has examined different types of funding scenarios for EOSC Exchange.

- One funding scenario focuses on funding dematerialised (remote) services such as downloading online data or services

EOSC Association AISBL
Rue du Luxembourg 3, BE-1000 Brussels, Belgium
+32 2 537 73 18 | info@eosc.eu | www.eosc.eu
Reg. number: 0755 723 931 | VAT number: BE0755 723 931

- Another one requires researchers to travel and go to a specific place (e.g. physical access to a machine).

The questions are related to public procurement rules that the EOSC Association has to abide by. To be accurate in legal terms, we recommend that a specialist of public procurement gives a written analysis.

Then, we will have to highlight different cases: who will ask for funding for the services? A given service provider? The association? The marketplace? Other?

Different funding scenarios integrate the nature of the services. Therefore, the answer to the following questions provides a guide to the development of such scenarios:

1. What is the nature of the services?
    a. The provider provides a service that is dematerialised e.g. remote access to data from a database as an example.
    b. The provider provides a service such as physical access to a machine which requires travel (for physical cross-border services) – these questions help us anticipate the VAT issue.
2. What is the provision of a service against payment?
    a. Who will invoice?
        i. With VAT?
        ii. Without apparent VAT? The institution(s) as beneficiary(ies) of the service collect, pay and deduce VAT to its own administration?
    b. Who will receive the invoice
        i. Each service beneficiary?
        ii. EOSC Association having a mandate?
        iii. The EOSC marketplace holder?
        iv. Other with a mandate?
    c. Who will pay?
        i. Member states: A dedicated budget line for the Ministry for the service providers? For the beneficiaries?
        ii. Beneficiaries?
        iii. Other?

EOSC Association AISBL
Rue du Luxembourg 3, BE-1000 Brussels, Belgium
+32 2 537 73 18 | info@eosc.eu | www.eosc.eu
Reg. number: 0755 723 931 | VAT number: BE0755 723 931

Appendix C - Case studies of data federations

C.1 Blue Cloud

The European landscape of **marine and ocean data management** has greatly advanced during the last three decades to develop standards and services, and to establish dedicated research and data infrastructures that cover the whole data value chain. Marine and ocean data are instrumental for the research, monitoring, evolution and management of the marine environment, but also for assessing fish stocks and biodiversity, supporting offshore engineering, hazard and disaster management, tourist industry, and many other socio-economic activities at sea and along the coasts. EU institutions employ the data to implement legislation, such as the Marine Strategy (MSFD)[61] and Water Framework Directives (WFD)[62]. To make the most of all these efforts, and to enable the use of cloud-based open science for ocean sustainability, the Blue-Cloud project[63], part of the H2020 'The Future of Seas and Oceans Flagship Initiative', has developed a thematic marine cloud that serves the interests of the Blue Economy, Marine Environment and Marine Knowledge agendas. At the base of the marine science "ecosystem" are the 'operators of instruments' and data collectors. On top of them sit the 'aggregators of data / service providers', such as COPERNICUS Marine Service[64], EMODnet[65] and SeaDataNet[66], who undertake activities for structured data management activities such as validation, standardisation, long term stewardship, and wider distribution. The system relies on the principle of "capture once – use many times" and on the achievement of 'data FAIRness' to ensure that maximum benefit can be derived from data once acquired.

Blue Data Infrastructures and DD&AS: Marine data are mostly collected by European marine data management infrastructures known as Blue Data Infrastructures ("BDIs"). These are developed and operated by research, governmental, and industry organisations from MS/AC and form the first element of the data value chain in marine science; they manage metadata, datasets, best practices, and associated services for import, discovery, and delivery on behalf of their networks of data-collecting and/or -managing organisations. By establishing links to data originators, BDIs engage and oversee the complete process from data collection to validation to storage and distribution, and can become involved in generating data products and models.

In order to facilitate users the discovery, retrieval and analysis of datasets and other data products, Blue-Cloud has implemented the **Data Discovery & Access Service (DD&AS) platform**[67], which constitutes a data federation of multidisciplinary repositories, analytical

[61] https://bit.ly/EU-MSFD
[62] https://ec.europa.eu/environment/water/water-framework/index_en.html
[63] https://blue-cloud.org/. The project started in October 2019 and will last until March 2023.
[64] https://marine.copernicus.eu/
[65] https://www.emodnet.eu
[66] https://www.seadatanet.org/
[67] https://blue-cloud.org/services/blue-cloud-data-discovery-and-access-service

EOSC Association AISBL
Rue du Luxembourg 3, BE-1000 Brussels, Belgium
+32 2 537 73 18 | info@eosc.eu | www.eosc.eu
Reg. number: 0755 723 931 | VAT number: BE0755 723 931

tools, and computing facilities from the BDIs. Currently, the Blue-Cloud DD&AS gives access to over **10 million data sets** as managed and provided by the federated BDIs. The DD&AS platform, as an essential tool to facilitate collaborative research in marine science and beyond, constitutes a strong support for EU's Green Deal and UN's Agenda 2030.

Data Services (in-kind from MS): MS provide many in-kind data services through their national activities as deployed by research institutes and national bodies. Data collection is mostly done at national level, as part of scientific activities or of governmental monitoring for environmental management or safety purposes. On a second level, MS deploy monitoring programs at national and regional level to support national marine governance goals, EU policies, and growing demands for sustainable use of resources. This data are complemented by further data gathered by research institutes through e.g. research vessels, and furthermore by data collected by private organisations in support of economic activities. The accompanying data management is also done mostly at national level[68]. MS/AC are also heavily involved in the implementation of EU directives that rely on the data to determine the current environmental status of their marine ecosystems, identify threats and establish measures to restore or maintain ecosystem health.

Current Sources of Funding: In kind contributions from MS/AC coexist with funding by the EU for *exchange systems* and common standards at European level that allow overview of and access to the data collections (this is the most visible part of the funding landscape; however, in a situation that resembles that of an iceberg, public and private funding at MS level underneath is actually much larger). Furthermore, most BDIs have been developed and built 'bottom-up' from community initiatives with co-funding from EU DG RTD and the ESFRI[69] programme, complemented since 2008 by funding from other EU Directorates, launching the 'top-down' initiatives EMODnet by EU DG MARE, COPERNICUS by EU DG DEFIS, as well as MSFD-related data projects by EU DG ENV and EEA—now further complemented by EOSC, which has received funding from DGs RTD and CNECT.

Participation of BDIs in these initiatives has benefitted them by further developing their sustainability base and helping propagate standards and services. This shows a clear need and opportunity for cooperation and synergy between RIs (i.e. BDIs) and the EU-lead initiatives for the benefit of both sides: RIs get access to networks of data originators and can manage pipelines for structured sharing of their data, while top-down initiatives focused on deriving data products and services for the community are thus able to obtain data from in-situ and remote sensing resources. For the Blue-Cloud DD&AS it is instrumental to align its position

[68] The contribution to data collection by international initiatives, like e.g. Copernicus, EU's Earth observation and monitoring programme (see https://www.copernicus.eu/en), cannot be neglected. Copernicus entered into force in 2014 and produces a wealth of data and information regarding the Earth sub-systems (land, atmosphere, oceans and inland waters) and cross-cutting processes (climate change, disaster management and security) which are mainly made available on a free, open and full basis.

[69] https://www.esfri.eu/

EOSC Association AISBL
Rue du Luxembourg 3, BE-1000 Brussels, Belgium
+32 2 537 73 18 | info@eosc.eu | www.eosc.eu
Reg. number: 0755 723 931 | VAT number: BE0755 723 931

with the European infrastructures and to ensure that the BDIs and major initiatives as EMODnet, COPERNICUS, and EOSC receive added-value.

C.2 CESSDA ERIC

The CESSDA ERIC (Consortium of Social Science Data Archives, European Research Infrastructure Consortium) is currently composed of 22 member countries and one observer. The mission of CESSDA ERIC is to provide a full-scale sustainable research infrastructure that enables the research community to conduct high-quality research in the social sciences contributing to the production of effective solutions to the major challenges facing society today. Member countries seek to increase the scientific excellence and efficacy of European research in the social sciences, as well as to expand easy access to data and metadata regardless of borders. They want to provide a research infrastructure for their researchers and join forces among their (national) data service providers.

CESSDA ERIC relies on the MS funding for its core functioning. Coordination, management, and administration, as well as a number of tools, services, and any internal developments' costs related to the whole consortium are covered by the core consortium budget. For example, the CESSDA Data Catalogue (CDC), which is the mandatory metadata aggregator for all service providers, is developed by contributions of all service providers and is maintained by CESSDA ERIC. Additionally, in each CESSDA member country there is a designated service provider (data archive or service). CESSDA service providers are usually part of public institutions (universities, research performing organisations, academies, foundations and data archives). The services they provide in their country are financed on the national level. The core development of CESSDA requires alignment of national legislation, policies, and procedures.

The CESSDA community relies on EC funding for contribution to the European SSH and cross-domain initiatives, the broader RI landscape in Europe, and connection to EOSC. CESSDA coordinates the SSHOC cluster[70] project (one of 5 INFRAEOSC-04 call funded projects). In this setting CESSDA and its partners develop new tools and service to connect the SSH community to EOSC facilitating access to flexible and scalable research data and related services. CESSDA increases data re-use through open science practices and FAIR principles to data management, and plays a major role in setting up a governance model for the SSH area of the EOSC. CESSDA fosters synergies across SSH disciplines by connecting established and new emerging communities. EC funding also covers cross-domain collaborations (BY-COVID[71] project), or connection to thematic SSH communities (HumMingBird[72], or COORDINATE[73] projects). CESSDA participates in projects like TRIPLE[74] or EOSC Enhance[75].

[70] https://www.sshopencloud.eu/
[71] https://elixir-europe.org/news/covid-new-eu-project-coordinated-elixir
[72] https://hummingbird-h2020.eu
[73] https://www.coordinate-network.eu/
[74] https://project.gotriple.eu/
[75] https://eosc-portal.eu/enhance

EOSC Association AISBL
Rue du Luxembourg 3, BE-1000 Brussels, Belgium
+32 2 537 73 18 | info@eosc.eu | www.eosc.eu
Reg. number: 0755 723 931 | VAT number: BE0755 723 931

C.3 Biodiversity Information community towards DiSSCo ERIC

DiSSCo RI (Distributed System of Scientific Collections Research Infrastructure)[76] works for the digital unification of all European natural science assets under common curation, access policies and practices that aim to make the data easily Findable, more Accessible, Interoperable and Reusable (FAIR). This infrastructure aims to overcome the existing fragmentation of Natural Sciences Collections (NSC) in Europe, building on existing databases, services and standards, by federating them. DiSSCo is currently supported by 24 countries (including Switzerland and Israel after March 2022 GA) and is now in its preparatory phase towards becoming an ERIC.

In its current phase, the in-cash and in-kind contributions from the coordinating country and institution (the Netherlands and Naturalis, respectively) are used to maintain the Coordinating Supporting Office, plus membership fees from voting members of the interim General Assembly, who participate either as national consortia or individually[77]. Additionally, DiSSCo as an infrastructure is based on the success of four key pillars in form of international actors:

- *Catalogue of Life (CoL*), a community-endorsed, agreed trusted source of reusable, standardised data for research in taxonomy and other disciplines, supported by (mostly public) institutions;
- *Taxonomic Databases Working Group (TDWG, today known as Biodiversity Information Standards)*[78], the reference for data exchange standards between users in many fields of bio- and geo-diversity;
- *Global Biodiversity Information Facility (GBIF)*[79], making data and standards, best practices and open-source tools accessible to anyone *via* the internet using web services;
- *Consortium of European Taxonomic Facilities (CETAF)*[80], network organisation providing a platform for its members to promote collection-based research into biodiversity and geodiversity, training in systematic biosciences and paleobiology

DiSSCo is funded by MS through projects (e.g. Mobilise COST Action, with funding from the European Cooperation in Science and Technology (ECST), i.e. not directly from the EC) and DiSSCo-linked projects to ensure its implementation (e.g. SYNTHESYS+ and DiSSCo-Prepare with funding from Horizon 2020). DiSSCo-related or -specific national or regional projects are used for e.g. mass digitization or digitalization on demand of the collections, acquisition of

[76] https://www.dissco.eu/

[77] This funding model will be revised after DiSSCo becomes an ERIC in 2024, when MS will become GA representatives.

[78] Incorporated in California (USA) with non-profit status, is funded mainly by individual (US$75) and institutional (US$500) membership fees, plus income from its annual conference

[79] Member organisation, with governments or organisations (both national, international or intergovernmental) as members.

[80] AISBL with 71 institutional members from 22 European countries. Until recently, they were mainly financed through membership fees.

EOSC Association AISBL
Rue du Luxembourg 3, BE-1000 Brussels, Belgium
+32 2 537 73 18 | info@eosc.eu | www.eosc.eu
Reg. number: 0755 723 931 | VAT number: BE0755 723 931

equipment for digitization, or to update the collections to the maturity level required. National contributions to DiSSCo's implementation were estimated at over 80 million € in total for 2021, mainly with funding from local roadmap projects. DiSSCo participates in the EU-funded project Biodiversity Community Integrated Knowledge Library (BiCiKL) together with major RIs in the domain, which looks at these synergies in practice, including collaborations and visibility of the communities' services within EOSC.

C.4 Covid-19 Data Platform

The Covid-19 Data Platform has been set up as a joint effort by the European Commission and EMBL's European Bioinformatics Institute (EMBL-EBI), together with EU MS and research partners such as ELIXIR. It addresses the need to share Covid-19 related data and findings in a coordinated way, to understand the disease and to develop treatments and vaccines at the start of the pandemic in 2020 when no central repository existed for this information. The Data Platform enables the rapid collection and comprehensive data sharing of available research data from different sources making it readily available to the European and global research communities, as well as comprehensive analyses using tools developed by EMBL-EBI and the wider scientific community.

The European COVID-19 Data Platform consists of three connected components:

- **SARS-CoV-2 Data Hubs**, which organise the flow of SARS-CoV-2 outbreak sequence data and provide comprehensive open data sharing as well as deep user support for the European and global research communities including public health agencies and hospital and academic laboratories
- **Federated European Genome-phenome Archive**, which provides secure controlled access sharing of sensitive patient and research subject data sets relating to COVID-19, e.g. omics from human research subjects, sensitive clinical and epidemiological data. Such data are often restricted within national borders and the federated EGA offers a national database with international connectivity
- **COVID-19 Data Portal**, which brings together and continuously updates relevant COVID-19 datasets and tools, hosts sequence data sharing and facilitates access to other SARS-CoV-2 resources. The European COVID-19 Data Portal includes a federation of national data portals, hosted in those nations.

There are more than 15 database resources[81] which actively feeds into the COVID-19 Data Platform, many of which are EMBL-EBI's data resources, including ENA, UniProt, PDBe, EMDB, Expression Atlas and Europe PMC. It is part of EMBL's and thus also EMBL-EBI's mandate to provide long-lasting data resources to the life science community. Since EMBL is itself funded by MS, this means that the federated data resources do rely, at least to some extent, on MS funding. However, it is important to note that EMBL-EBI's data resources also receive significant co-funding from national public and private funding organisations from Europe and

[81] https://www.covid19dataportal.org/database-resources

EOSC Association AISBL
Rue du Luxembourg 3, BE-1000 Brussels, Belgium
+32 2 537 73 18 | info@eosc.eu | www.eosc.eu
Reg. number: 0755 723 931 | VAT number: BE0755 723 931

beyond (e.g. NIH, Sanger, Wellcome Trust, etc.). Other database resources not hosted at EMBL-EBI include the Human Protein Atlas, funded by the Knut and Alice Wallenberg Foundation in Sweden. Furthermore, national funding was provided to set up the respective national Data Portals.

EC funding from various projects was used to set up and make operational the different Data Platform components i.e. the Covid-19 Data Portal (which included setting up a template for the national portals), and SARS-CoV-2 Data Hubs, as well as to develop all computational and data workflows and mobilisation necessary to effectively synchronise different data resources with the Platform. It is worth noting that project money was also provided to fund legal expertise which was complemented by in-kind contributions from other partners.

EOSC Association AISBL
Rue du Luxembourg 3, BE-1000 Brussels, Belgium
+32 2 537 73 18 | info@eosc.eu | www.eosc.eu
Reg. number: 0755 723 931 | VAT number: BE0755 723 931