EOSC-hub Briefing Paper – Provision of Cross-Border Services

Status	Final
Version	2.1
Delivery Date	16 October 2020
Dissemination Level	Public

About this paper

This Briefing Paper results from work conducted jointly by EOSC-hub T2.3 and WP12 examining EOSC use cases involving cross-border consumption of rivalrous resources and analysing the EOSC using an approach to business modelling called Platform Design. The paper discusses challenges involved in provision of publicly funded services such as storage and compute across borders, identifies key EOSC participant types and the opportunities the EOSC could present for them, and presents recommendations for the EOSC. The paper was the subject of a public consultation and small updates were made as a result of feedback received.

EOSC-HUB [illegible] FUNDING FROM THE EUROPEAN UNION'S HORIZON 2020 RESEARCH AND INNOVATION PROGRAMME UNDER GR[illegible]EMENT NO. 777536.

Delivery slip

From	**Editor:** Dale Robertson, Jisc **Core Team/Authors:** Sergio Andreozzi (EGI), Magchiel Bijsterbosch (SURF), Luciano Gaido (INFN), Antti Pursula (CSC), Dale Robertson (Jisc) **Contributors/Acknowledgements:** Owen Appleton (EGI), Giacinto Donvito (INFN), Tiziana Ferrari (EGI), Nadina Foggetti (INFN), Matti Heikkurinen (EGI), Hanno Holties (ASTRON), Ari Lukkarinen (CSC), Ludek Matyska (CESNET), Bob Jones (CERN), Hylke Koors (SURF), Steven Newhouse (EBI-EMBL), Raymond Oonk (SURF), Roberto Pizzo (ASTRON), Paul Rouse (GÉANT), Miroslav Ruda (CESNET), Luca Ruggeri (Boundaryless), Davide Salomoni (INFN), Gergely Sipos (EGI), Mark van de Sanden (SURF), Olaf Verschoor (GÉANT), Michelle Williams (GÉANT) **Approved by:** EOSC-hub Project Management Board

Acknowledgement

The team would like to acknowledge the contribution of Boundaryless, providers of the Platform Design Toolkit, who facilitated the discussions applying the platform design methodology to the EOSC. Your highly professional and very valuable guidance and expertise were very much appreciated!

Contents

1. Executive Summary

This paper describes challenges involved in provision of publicly funded rivalrous resources[1] across borders in Europe, with a focus on demanding use cases involving large amounts of resources. It discusses opportunities the EOSC could present for key participants and puts forward recommendations for the EOSC. The paper complements and draws on EOSC-hub deliverables D12.1[2] and 12.2[3], the EOSC-hub proposals for the EOSC Federating Core[4] and the working draft Community Position Paper on the EOSC Federating Core[5]. A public consultation was held which provided supportive feedback but clarified some points around the expected governance of the EOSC, as a result of which small updates were made to the recommendations to align them with the expected distribution of governance responsibilities. The recommendations are aimed primarily at the EOSC Working Groups and governing bodies, national and European e-Infrastructures and Research Infrastructures, researchers and Research Producing Organisations.

The research data cycle requires data, storage, networking and computing and the EOSC business model needs to address all of these aspects but the work was motivated by the observation in the draft Community Position Paper that the EOSC should align investments from member states with the compensation of marginal costs associated with cross-border usage of depletable resources and services. The paper thus focuses on how Shared Resources could be provided across borders. The Shared Resources include rivalrous resources such as storage and compute services which are often provided in the research environment by national publicly funded e-Infrastructures.

The basis for the work is that excellent research in the modern age requires access to high-quality data and world-class infrastructures and other resources such as software. The European research landscape is very fragmented with a large number of actors involved, often providing or consuming services within the boundaries of their thematic areas and national borders.

To analyse in detail the issues involved in research collaboration where cross-border storage and compute services are needed, two use cases were conducted, simulating negotiations between publicly funded service providers (CESNET, CSC, INFN and SURF) on the supply side and the ELIXIR community and the Low Frequency Array (LOFAR) telescope experiment on the demand side. The ELIXIR use case involved investigating outsourcing of the computing platform for analysis of ELIXIR datasets (of the order of petabytes in size) by hundreds of research groups, in combination with their

[1] Rivalrous - the consumption/use of the resource by one person/organisation reduces the availability of the resource to another person/organisation. Storage and compute services fall into this category and are the focus of the work in this paper.

[2] https://www.eosc-hub.eu/deliverable/d121-procurement-requirements-and-demand-assessment, June 2019

[3] Draft available for consultation during August 2020, see https://www.eosc-hub.eu/news/participate-eosc-hub-procurement-and-business-models-consultation

[4] https://www.eosc-hub.eu/sites/default/files/EOSC-hub%20Briefing%20Paper%20-%20EOSC%20Federating%20Core%20Governance%20and%20Sustainability%20v1.0_0.pdf and https://www.eosc-hub.eu/sites/default/files/EOSC-hub%20Briefing%20Paper%20v2.0%20-%20EOSC%20Federating%20Core%20v0.3%20%28consultation%20comments%20and%20responses%29%20%282%29.pdf

[5] See https://www.eosc-hub.eu/publications/eosc-federating-core-community-position-paper-v11

own data. The ASTRON use case concerned requirements for offering a data processing service to the LOFAR community to help the generation of science-ready data products from data stored in the "instrument" data archive, which is distributed over federated sites in Holland, Germany and Poland. For each use case, representatives of the users and providers met together (virtually) for sessions dedicated to discussion of the technical aspects and the commercial aspects, followed by one or more sessions to assess the results and identify conclusions and lessons learned.

In parallel, a business modelling open source methodology for platform ecosystems, called the Platform Design Toolkit[6], was deployed to model the use cases with the aim of identifying ways in which the EOSC could facilitate interactions between users and service providers. The PDT approach models a set of entities which need to or want to interact and assists them in doing so by shaping a role for a platform (possibly supported by the appropriate technology) to provide the means for producers and consumers (service providers and users, in the case of the EOSC) to connect. The aim is to define a strategy that creates network effects, in which the "ecosystem" evolves and grows, increasing in value as more entities use it.

The use case and platform design work confirmed that the EOSC has the opportunity to reduce barriers and enable frictionless, easy access to data and related services across geographic boundaries and disciplines so research communities can better connect with suppliers, users and funders, and to promote a cross-border multidisciplinary environment where past, present and future investments in resources and infrastructures can be efficiently leveraged and benefit from economies of scale. Numerous enabling services and activities are, however, required to achieve this.

The use cases highlighted that for demanding or high-volume cases, complex information needs to be accessed and exchanged before transactions (i.e. contracts for the supply of publicly funded compute and storage resources) can be concluded. The EOSC can provide the necessary framework to put demand and supply in contact efficiently and consistently but to do so the technical and commercial information it provides needs to go beyond what is currently being implemented

The majority of research funding in Europe is provided nationally. Funding sources are varied, complex and involve a large number of different rules, which contributes to suboptimal use of member states' investment in research services, particularly in cases of cross-border service usage. Research use cases could be aided by providing "choreography" of national and European funding schemes and eligibility criteria across borders.

Insight 1:

Rapid transactions are a key to the success of the EOSC. Capabilities are required of the EOSC as a platform, to support negotiation and matchmaking for complex and demanding use cases, including:

- standardised and optimised processes, regulations and agreements involved in transactions matching supply and demand

[6] See https://platformdesigntoolkit.com/

- services listings in more numerous, standardised and granular categories which permit different forms of value-added support to be distinguished (e.g. commodity services; added-value services such as consultancy, collaboration, commodity resources including support; etc)
- demand-side advertising of needs and challenges
- support for the use of supplier reputation, as a driver of quality
- a knowledge base to signpost information about national funding sources.

Recommendation 1:

The EOSC has the opportunity to make matchmaking a standard process with a high rate of success. It should provide the capability to put demand and supply in contact efficiently and consistently, including for complex use cases, and facilitate negotiation, by acting as a platform to signpost information, standardise and support dialogue and optimise transactions.

The role of Research Support Providers[7] throughout the process of matching supply and demand is key to making use of information, to easing the effort and difficulty involved and to supporting the long tail of users to collaborate more strongly in research communities: the complexity and individual nature of many of the transactions which need to be encouraged mean that the EOSC as an information source, however rich, is not sufficient to deliver the "mass personalisation" potential of the platform. The knowledge and expertise of qualified, experienced Research Support Providers is needed to successfully match demand to supply, reducing friction in transactions. Research Support Providers would be amongst the most intensive users of the EOSC, challenged with becoming familiar with its offers or demands (as appropriate), and matching them up in the most effective way.

Insight 2:

Qualified, experienced Research Support Providers – providers of expert technical support and consultancy - are key to reducing friction in research transactions. Their input is required to successfully boost research collaboration and make more effective use of research assets, for example to

- help users to understand the implications of different choices
- build demonstrators and assess technologies
- help arrange use of services whilst protecting users from lock-in and supporting movement between providers.

Recommendation 2:

Create an EOSC Coordination Point providing a federated European network of distributed expert teams, mobilising existing Research Support Provider resources, using resources from national e-Infrastructures (funded nationally) as a basis, and provide a small EOSC coordinating team funded by the EOSC, for example as part of the EOSC Secretariat.

[7] For example, support staff at Research Infrastructures, e-Infrastructures or universities

The capability to organise the exploitation of large-scale distributed datasets requires compute and storage resources to be adequately provisioned and integrated. Funding their cross-border use is complex and challenging, but to support international research collaboration a means needs to be found of doing this in a reliable manner which involves a proportionate amount of effort for research communities. Existing mechanisms such as Virtual Access, procurement frameworks and vouchers, whilst they each have their merits, do not necessarily provide suitable solutions.

It is challenging to coordinate research funding to meet the needs of international research collaborations for services supported by depletable capacity, such as storage and compute, particularly for multidisciplinary research collaborations. Shared Resources not only need to be sustained, but also a way needs to be found for international research communities to pay service providers which have national scopes and remits. A more frictionless marketplace needs to be created, in which funds flow to those service providers which deliver the most value. Current EC or international funding schemes are not well suited to funding cross-border, cross-disciplinary consumption of depletable resources such as storage and compute. There is at present no straightforward means for a research group to pay with EC research funding (for example European Research Council funding) for use of existing e-Infrastructure, but it is possible for such funding to be used to build or develop new infrastructure. This increases fragmentation of digital service provision and reduces utilisation of existing e-Infrastructure, highlighting the suboptimal use of funding. There needs to be a tighter coupling between research funding and e-Infrastructure funding, at the EU level and also on a national level.

Shared resources such as storage and compute are a fundamental part of the EOSC: as is the case with high-speed networking, there is a need to provide such services outside of the researcher's own domain. The EOSC Partnership needs to ensure appropriate coordination and synchronisation of member state and EC funding to sustain the whole EOSC ecosystem including the Shared Resources. The mechanics for funding the shared resources are complex. Centrally held European money is difficult to allocate but has the advantage of being somewhat blind to geography and some EC co-funding of access to national infrastructures may provide the necessary "glue". On the other hand, nationally held funds or in-kind contributions may be easier to administer and should be considered for inclusion as part of the minimal viable EOSC.

Insight 3:

Shared Resources such as storage and compute are a fundamental part of the value-add of the EOSC. A means needs to be found to fund their cross-border use in support of international research collaboration. Measures which may help include:

- studies to obtain further empirical evidence of the macroeconomic and societal benefits of cross-border research funding
- support of cross-border research collaborations organised in pilots to help assess the potential value of data and other resources for new use cases

- developing guidelines and structures that facilitate cross-border public-to-public service collaboration with cost reimbursement (e.g. public-to-public cooperation exemption from public procurement rules)
- providing supplementary demand-side funding for use of national e-Infrastructures, including their cross-border use, to complement or top-up national supply-side funding
- payment of member state contributions to the EOSC Partnership partly in the form of in-kind donation of resources by member states to the EOSC, to create a tighter coupling between national e-Infrastructures and the EOSC.

Recommendation 3:

The EOSC Partnership needs to ensure appropriate coordination and synchronisation of different funding streams (e.g. from member states, EC, research performing organisations and projects etc) to sustain the whole EOSC ecosystem including the Shared Resources. To underpin this policy, dialogues should take place within member states between ministries, research funders and Research Service Providers to allow for a cohesive input and contribution across the operational support of pan-European research, representation in the EOSC Partnership, and the broader EU research policy.

The work has demonstrated ways in which the EOSC can develop to meet its objectives but it has also highlighted the suboptimal use of public member state funding: long-term strategies on science and infrastructure are not necessarily formulated in a joined-up manner, and EC and member state funding do not interoperate efficiently. The majority of e-Infrastructure investment is still funded nationally. It is challenging to coordinate it to meet the needs of an international research collaboration. On a thematic level coordinated activity does take place, but it is particularly challenging to coordinate funding for e-Infrastructures services in support of multidisciplinary research collaboration. The complexity and fragmentation of funding rules for cross-border research activities is a recognised issue highlighted by the EOSC, which needs to be addressed, but it has received limited attention during the implementation of the EOSC to date. The insights described in this paper and the conclusions in D12.2, point the way to possible options for funding and providing Shared Resources across borders, which should be investigated further.

The EOSC would not fulfil its mission just by organising a web of FAIR data; the EOSC has clear potential to act as a platform stimulating further research collaboration, producing more and better science to the benefit of all, but it involves the practical implementation steps recommended here, in particular accepting the cost of sustaining Shared Resources, the demand for which has been illustrated by the use cases. During the consultation on the paper's proposals, one correspondent observed that the Shared Resources are needed in the EOSC *from the start,* because [access to] data on its own is not enough. Many of the necessary services and consultancy resources are in place, although they operate mainly within national or community boundaries and the importance of their sustainability beyond the service development phase needs to be better recognised and supported. Research Support Providers, particularly those playing an established role in the national and European e-Infrastructures, also have a crucial role to play, acting as trusted advisors. The basis for the platform already exists within the research ecosystem and can be built upon.

The demand exists for cross-border use of assets. To allow more and better science to be performed, there is a clear incentive for governments and national funders to facilitate cross-border use of assets. The e-Infrastructures are recognised as the horizontal foundation of the EOSC; the work described here confirms this view.

2. Introduction

This Briefing Paper describes challenges involved in provision of publicly funded rivalrous resources[8] such as storage and compute across borders in Europe, with a focus on demanding use cases involving large amounts of resources. It discusses opportunities the EOSC could present for key participants and puts forward recommendations for the EOSC, on which feedback is invited. It is the outcome of work to examine EOSC business models conducted jointly by a team from EOSC-hub T2.3 (Governance and Sustainability) and WP12 (Business Models and Procurement) including representatives from EGI, EUDAT, INFN, Jisc and SURF, between January and July 2020. The paper's proposals were the subject of a public consultation from 14 September to 9 October 2020, supported by a webinar on 15 September 2020, and were also discussed in the EOSC Architecture Working Group on 16 September 2020. The feedback received was supportive but clarified some points around the expected governance of the EOSC, as a result of which small amendments were made in this paper, including to its recommendations.

The basis for the work is that excellent research in the modern age requires access to high-quality data and world-class infrastructures and other resources such as software. The European research landscape is very fragmented with a large number of actors involved, often providing or consuming services within the boundaries of their thematic areas and national borders. The EOSC has the opportunity to reduce barriers and enable frictionless, easy access to data and related services across geographic boundaries and disciplines so research communities can better connect with suppliers, users and funders, and to promote a cross-border multidisciplinary environment where past, present and future investments in resources and infrastructures can be efficiently leveraged and benefit from economies of scale.

The paper builds on the EOSC-hub proposals for the EOSC Federating Core, the working draft Community Position Paper on the EOSC Federating Core, and deliverables D12.1 and 12.2. The Federating Core proposals were shared for consultation during 2019[9]. The consultation, in particular discussion with the thematic clusters[10], produced a vision in which the unique value-add of the EOSC is its ability to provide quality assessment, combined with the ability for researchers to access and reuse data alongside Shared Resources[11] through the same portal. The Community Position Paper proposes that "the EOSC needs to sustain the costs of providing the benefits of open data policies to a wider community of users. It needs to create the financial vehicle to cover the costs of…provision and consumption of the Shared Resources beyond their originating communities. Coordinated

[8] Rivalrous - the consumption/use of the resource by one person/organisation reduces the availability of the resource to another person/organisation

[9] https://www.eosc-hub.eu/sites/default/files/EOSC-hub%20Briefing%20Paper%20-%20EOSC%20Federating%20Core%20Governance%20and%20Sustainability%20v1.0_0.pdf and https://www.eosc-hub.eu/sites/default/files/EOSC-hub%20Briefing%20Paper%20v2.0%20-%20EOSC%20Federating%20Core%20v0.3%20%28consultation%20comments%20and%20responses%29%20%282%29.pdf

[10] https://www.eosc-hub.eu/eosc-hub-and-esfri-cluster-projects

[11] Resources including scientific outputs (local copies of data; applications, software, pipelines etc.) and the storage and compute hosting platforms needed to deposit, share and process them

provisioning and funding of the Federating Core is expected to bring economies of scale by aligning investments from member states with the compensation of marginal costs associated with cross-border usage of depletable resources and services". The Shared Resources include rivalrous resources such as storage and compute services which are often provided in the research environment by national publicly funded e-Infrastructures. A summary of the EOSC-hub proposals for the organisation and contents of the Federating Core is included in Appendix I.

D12.1, Procurement Requirements and Demand Assessment, performed a demand-side analysis of digital services for research, and proposed possible models for acquiring digital services in the EOSC context. D12.2, Report on Business Model Analysis for Procuring Services in the EOSC, further develops the analysis of the issues involved in acquiring publicly funded digital services for research in the context of the EOSC and explores possible solutions based on case studies. This Briefing Paper complements and draws on D12.1 and D12.2.

The work has been taking place in the wider context of a gathering pace of activity to implement the EOSC. The EOSC-hub project[12] contributes to several of the strategic action lines involved in implementing the European Open Science Cloud[13], as do numerous other Horizon2020-funded projects[14]. Work to establish the second phase of the EOSC from the start of 2021 has produced a draft EOSC Partnership Agreement document[15] which proposes the continuation of the EOSC as a Co-Programmed Partnership of Horizon Europe and defines strategic objectives and outline KPIs for the EOSC. The EOSC Strategic Research and Innovation Agenda (SRIA) document was in public consultation until 31 August 2020[16]. Once finalised, it will be a key document in defining the EOSC and will feed into the work programmes which will define future projects and funding under Horizon Europe[17]. In parallel to this, the definition and formation of the EOSC governance bodies for 2021 onwards is moving forward although this is still in progress. The INFRAEOSC-03 and -07 funding calls[18] closed in June and the projects to be funded as a result are currently in preparation. The projects the calls will fund will contribute to EOSC implementation activity, including operation of the EOSC-Core, until mid-2023[19]. The EOSC Executive Workplan 2020[20], published in May 2020, includes details of the planned activities of the EOSC Executive Board and its Working Groups. The EOSC Executive will present a minimal viable EOSC at the formal EOSC launch event in December 2020. The minimal viable EOSC is being defined by ongoing work in the EOSC Working Groups[21]. The EOSC-hub Federating Core

[12] https://eosc-hub.eu/

[13] Contained in https://ec.europa.eu/research/openscience/pdf/swd_2018_83_f1_staff_working_paper_en.pdf

[14] For a list of Horizon2020 projects relevant to the EOSC, see the Annex of the EOSC Strategic Implementation Plan available from https://www.eoscsecretariat.eu/eosc-governance/eosc-executive-board

[15] https://ec.europa.eu/info/files/european-open-science-cloud-eosc_en

[16] https://www.eoscsecretariat.eu/open-consultation-eosc-strategic-research-and-innovation-agenda

[17] https://ec.europa.eu/info/horizon-europe-next-research-and-innovation-framework-programme_en

[18] See https://www.eosc-portal.eu/infraeosc-03-2020-integration-and-consolidation-existing-pan-european-access-mechanism-public and https://www.eosc-portal.eu/infraeosc-07-2020-increasing-service-offer-eosc-portal

[19] See https://ec.europa.eu/digital-single-market/en/news/proposals-scaling-european-open-science-cloud, published 13 July 2020

[20] https://op.europa.eu/en/web/eu-law-and-publications/publication-detail/-/publication/ae215698-af7b-11ea-bb7a-01aa75ed71a1

[21] https://www.eoscsecretariat.eu/eosc-working-groups

proposals provided input to this process (information on the contributions made is included in Appendix I) and the consultation version of this Briefing Paper is referenced in the Sustainability Working Group's draft Iron Lady document[22] which iterates from the its earlier Tinman document[23].

The work consisted of three main phases: (1) two simulated negotiations between research organisations representing the demand side and publicly funded providers on the supply side (see Chapter 3: Use Cases); (2) applying platform design thinking to the EOSC (see Chapter 4: EOSC Ecosystem); (3) analysing insights drawn from these activities and from D12.1 and D12.2, combined with information from a session on Issues in Cross-Border Consumption of Resources[24] and formulating recommendations (see Chapter 5: Insights and Recommendations). Conclusions are drawn in Chapter 6. Appendix I contains a summary of the EOSC-hub Federating Core proposals and comparison with the Sustainability Working Group Tinman proposals. Appendix II contains a detailed description of the two use cases. Appendix III contains a full description of the platform design methodology applied to the EOSC.

The research data cycle requires data, storage, networking and computing and the EOSC business model needs to address all of these aspects but the work presented in this document was motivated, by the observation in the draft Community Position Paper that the EOSC should align investments from member states with the compensation of marginal costs associated with cross-border usage of depletable resources and services. The paper thus focuses on how Shared Resources could be provided across borders. The proposed recommendations are a key step on the EOSC-hub Governance and Sustainability Roadmap and complement the work of D12.1 and D12.2, but consultation with the wider community was also undertaken, to gather input on their suitability. The recommendations in this Briefing Paper are aimed primarily at the EOSC Working Groups and governing bodies, national and European e-Infrastructures and Research Infrastructures, researchers and Research Producing Organisations.

[22] See information at https://www.esfri.eu/sites/default/files/BobJones_Session-1_ESFRI-EOSC_06.10.20.pdf
[23] Available from https://www.e-ciencia.es/tinman-report/
[24] Held as part of the EOSC-hub week 18-20 May 2020. See https://www.eosc-hub.eu/eosc-hub-week-2020/agenda/issues-cross-border-consumption-resources-eosc

3. Use Cases

3.1 Description

For the simulated negotiation exercise two scientific communities representing good examples of the different needs foreseen within the EOSC ecosystem were selected: the ELIXIR[25] community, and the Low Frequency Array (LOFAR) telescope experiment[26]. Both use cases concerned provision and consumption of shared resources (specifically, storage and compute services) beyond their originating communities. While ideally the simulation would be repeated to create a larger sample size, the coinciding pandemic has prevented this within the timeframe available for this work. However, both use cases provide a valid illustration of the current practice and challenges that European Research Infrastructures face and a larger sample would perhaps add to but not change the picture that these two use cases paint.

The use cases' needs and demands, the offers from different service providers and the results of the negotiations are described in Appendix II. The main insights gained are described below.

Each use case involved (virtual) sessions between the user and the suppliers dedicated to discussion of the technical aspects and then the commercial aspects of the case, followed by one or two sessions to analyse and discuss the results and the issues identified during the simulations. The technical and commercial sessions were facilitated by mediators, i.e. experts involved in EOSC-hub T2.3 and WP12. The technical sessions discussed issues including the services to be operated, the amount of resources (computing and data) involved, authentication and authorisation, required levels of support, and compliance to standards. The commercial sessions discussed funding models and the legal and contractual requirements and challenges involved between the user and supplier, informing where EOSC efforts should best be directed to address the status quo.

Use case 1 concerned the ELIXIR community, represented during the simulation exercise by the European Bioinformatics Institute (EMBLl-EBI). The ELIXIR community consists of hundreds of small research groups 5-10 people each) across Europe. The scenario analysed was that of outsourcing the computing platform required by these research groups to analyse public ELIXIR datasets (of the order of Petabytes) in combination with their own private datasets, which may contain personal data. Each research project conducted by a group could potentially be executed with a different resource provider, providing that applications are portable across different cloud providers. However, for use cases which need small to medium amounts of computational and storage capacity for a short period of time, a combination of providers to meet the needs of a single project using a distributed model is not currently a common use case. Large projects involving many research groups have successfully used the distributed model when the extended period of their data processing and analytics activities and the amount of capacity requested have to be sustained through the federation of national investments and the use of existing digital facilities to reduce the overall cost of ownership. As the federation of providers requires processes for the coordinated delivery of services and the alignment

[25] ELIXIR is an intergovernmental organisation that brings together life science resources from across Europe: https://elixir-europe.org

[26] http://www.lofar.org

of common standards and policies, the federated model can be seen as complicated and costly, although this may become more common to avoid the movement of data across legal jurisdictions and/or make access to distributed large datasets technically and economically more efficient.

The simulation involved four service providers:

- CESNET (Czech Republic)[27]
- CSC (Finland)[28]
- INFN (Italy)[29]
- SURF (Netherlands)[30].

Use case 2 concerned the LOFAR Radio Astronomy Competence Centre, represented during the simulation by ASTRON, the Netherlands Institute for Radio Astronomy. ASTRON is responsible for the Low Frequency Array (LOFAR) telescope operations. It aims to offer a data processing service to the LOFAR community to help the generation of science-ready data products from data stored in the 'instrument' data archive. The latter is a distributed data archive (over 50PB) with federated sites in Amsterdam (Netherlands), Jülich (Germany), and Poznan (Poland), but more data centres might be involved in the future, depending on the interest of countries joining the International LOFAR Telescope. The simulation exercise involved three service providers:

- CSC (Finland)
- INFN (Italy)
- SURF (Netherlands).

SURF already supports ASTRON/LOFAR, while Finland and Italy are not participating in the LOFAR initiative for the time being.

The operational model is still under discussion and will likely evolve over time. The initial phase is envisioned to be a managed service where users apply for compute resources, and standard processing pipelines are run on behalf of the user for generating higher level data products from data in the instrument archive. Later phases may involve ASTRON providing a science data repository for science level data, and compute as a platform.

3.2 Discussion

The goal of the use cases was to simulate technical and commercial discussions between the demand side (a user community representative acting on behalf of various teams within the community) and the supply side (four providers for Use Case 1 and three for Use Case 2) to explore the opportunities for them to reach an agreement on the transnational provision of resources provided by publicly funded actors.

Although quite similar in their technical aspects, the two use cases differed with respect to funding: ELIXIR, through the institutions and projects with which it is engaged, is experimenting with use of a

[27] https://www.cesnet.cz/cesnet/?lang=en
[28] https://www.csc.fi/en/csc
[29] https://home.infn.it/en
[30] https://www.surf.nl/en

central pot of money so (theoretically) has money to spend and the main goal of the discussion was to determine the best provider(s) for their needs, while ASTRON has the additional issue of finding the money for the additional resources. One of the possible options to "pay" for the resources, is to establish a project collaboration (e.g. within an EC-funded project) between the research infrastructure (ELIXIR or ASTRON) and the service providers. For ELIXIR, this is one possible option available; for ASTRON this is the only option since they do not yet have a source of funding for their desired activity. All the service providers involved in the simulation exercise would potentially be interested in collaborating with ASTRON, but it is not an easy task and its exploration is time-consuming. It also depends heavily on funding streams at both European and national levels being available at the moment the need for resources arises. Although all the actors involved in the exercise have quite a lot of experience in this, it is still not necessarily straightforward and overall the use of a project collaboration as a way of funding the required services was seen as requiring a disproportionately high administrative and organisational burden, losing the advantage of the dynamic nature of providing compute/IT resources.

The use cases highlighted the impact of the European public procurement rules on public institutions' behaviour. ELIXIR is currently experimenting, for instance through the OCRE project[31], with how it might be able to allocate a central pot of money (i.e. money held and allocated by the ELIXIR Hub rather than by the in-country ELIXIR nodes[32]) in support of its Platform approach[33]. Aggregated procurement using a similar model to that for IaaS established by GÉANT and the NRENs[34] is, theoretically at least, a possibility; ELIXIR is continuing its consideration of this approach.

The ELIXIR simulation exercise showed that for the providers there can be a considerable difference to the offers should they be made directly from the institution (i.e. CESNET, CSC, INFN or SURF – and potentially also involving offers being made to provider services in a different country) or through the local national ELIXIR nodes. A request coming from the local ELIXIR node to a local provider may attract a different charging model, be part of an established resource-sharing agreement, or not trigger EU public procurement. CESNET, CSC and SURF all expressed their preference to provide services through their local ELIXIR node; by comparison even if ASTRON had funding for its use case, direct offers to provide services would need to be considered on a case-by-case basis taking into account national priorities due to procurement regulations. This demonstrates the caution of publicly funded service providers with respect to the public procurement rules, as a result of which suppliers mostly provide services only to organisations with in-house status (i.e. the paying organisation is part of the governance of the provider and can influence its activities); there is normally a limit of 20% for transactions on the open market, i.e. with organisations which do not have in-house status. On the other hand, INFN offered the possibility of providing services to ASTRON using a pay-per-use model by establishing a research agreement, on the basis that it is not necessary to run a public tender because of the combination of advanced IT services and specialised expertise for research purposes.

It was observed that neither of the use cases would be suited to vouchers or to Virtual Access. Vouchers are suitable for short-term, relatively small-scale usage but not for long-term large-scale

[31] https://www.ocre-project.eu/
[32] https://elixir-europe.org/about-us/who-we-are/nodes
[33] See https://elixir-europe.org/platforms/compute for example
[34] https://clouds.geant.org/geant-cloud-catalogue/geant-cloud-catalogue-iaas/

needs; Virtual Access, the mechanism currently used by the European Commission to pay for resources made available through EC-funded projects, is limited to services included in project proposals. Also, Virtual Access requires evidence of service usage, which makes it incompatible with services which do not require registration as it is not possible to detect new users beyond the local community, and so provide the necessary evidence to recover the cost of service usage. The nature of VA funding being based on actual usage rather than available (reserved) capacity can also present a risk to service providers which may invest in additional capacity to meet usage proposed through an EC-funded project, but where the usage does not materialise so expected income is never received. Being generally risk averse due to limited margins to cover these risks and the impact of public accountability on financial governance, publicly funded organisations will seek to avoid these kinds of demand-failure associated risks which may result in a net operating loss, thus resulting in a demand surplus. The use of vouchers is discussed in more detail in D12.2 which observes that the VAT considerations relating to vouchers create a significant overhead. Another consideration is that, although it is possible for at least some (for example INFN and CSC), if not all, publicly funded e-Infrastructures to invoice across borders, in practice this is difficult due to requirements around the calculation of costs, so providers would need to further develop their internal information and procedures to cater for this.

To allow for unlimited scalability with demand while leveraging economies of scale of existing public investments, ideally, there needs to be a cost recovery mechanism whereby the research infrastructure (e.g. ELIXIR or ASTRON/LOFAR) reimburses the costs of one or more of the publicly funded national e-Infrastructures in return for the provision of digital services. Discussion of the ASTRON use case led to the observation that supplementary budgets (e.g. grants to fund the infrastructure requirements of specific research) to co-fund use of national infrastructures would be helpful: this would provide additional demand-side funding to complement, or top up, the supply-side funding nationally. An example of this is the Dutch Fundamental Sciences e-Infrastructure (FUSE) programme[35] which has been funded under the National Roadmap for Large Scale Research Infrastructures.

One of the important elements that emerged during the ELIXIR simulation, is that while finding commodity resources (cpu, storage, etc) does not present major issues, it is quite difficult to understand which providers are able to provide value-added services on top of these. This is made more difficult because there are so many different levels of value-added services (ranging from commodity service plus support, to consultancy expertise up to services which are the result of the composition of basic services and collaborations) but the requirement for value-added services was expressed during the exercise. A first, minimal approach to overcome this could be to group services within the EOSC catalogue according to the different kinds of added value that can be provided along with the bare service.

The role of the facilitators during the technical and commercial discussions proved to be useful to better understand the requirements of the use cases and to support matchmaking with the service offerings. This was recognised as an added value with respect to what the EOSC is currently offering

[35] https://www.nwo.nl/en/research-and-results/programmes/nwo/national-road-map-large-scale-research-infrastructure/projects-2019-2020.html

through its catalogue and marketplace where, in complex and demanding use cases such as those illustrated here, it is quite difficult for an end user to really understand which services are suitable for his/her needs. Rather than a static process that is associated with, for example, an online shop, the dynamic interaction between demand requirements and supply capabilities and limitations is critical to come to a fit. It was observed that a distinction between public and commercial providers is that publicly funded e-Infrastructures typically provide such value-added services, highly specialised towards the needs of the research community to help them to generate research assets, whilst the commercial sector is generally less close to the community.

It was also noted during the exercise that the availability of high-speed networking capacity (local networks, NRENs and GÉANT) to connect the research data location to the required publicly funded storage and compute resources can generally be assumed, which is not necessarily the case for commercial services.

The use cases described in this chapter show that the process of acquiring research services to meet demand is non-standard: many use cases have technical, functional and/or business requirements which are highly specific or even unique. In the ELIXIR use case, the suppliers would only make offers through the local ELIXIR nodes, whereas for ASTRON the different suppliers were all prepared to consider working together in a funded project collaboration despite not being involved in a close prior partnership with LOFAR. If the EOSC were involved in bridging the gap between user and supplier, in the case of ELIXIR the EOSC would be involved locally – i.e. contracting and provisioning would be performed in each country individually - whereas with LOFAR the EOSC could be involved centrally, perhaps by helping to identify suitable EC project funding. ELIXIR, theoretically, had money to spend and was looking for a way to be able to spend it, whereas ASTRON needed to identify a source of money to implement their use case. Given the fragmentation and complexity the use cases illustrate, a portal focusing on promoting a standardised service catalogue is unlikely to be able to bring supply and demand together successfully.

4. EOSC Platform Design

4.1 Description

Complementing the two use cases described in chapter 4 was the use of the Platform Design Toolkit[36] (PDT), a business modelling open source methodology for platform ecosystems, to analyse and document consideration of EOSC business models, enabling insights to be captured. The methodology was selected from amongst several options[37] as the most suitable for modelling of the EOSC.

The PDT approach models a set of entities which need to or want to interact and assists them in doing so by shaping a role for a platform (possibly supported by the appropriate technology) to increase value exchange and derive new ways of working. The methodology involves studying the assets, challenges, goals and motivations of the entities and identifying opportunities for them to add or derive more value by interacting through a "platform". The platform would provide the means for producers and consumers (service providers and users, in the case of the EOSC) to connect. The aim is to define a strategy that creates network effects, in which the "ecosystem" evolves and grows, increasing in value as more entities use it. Uber and Airbnb are examples of platforms involving networking effects, which have opened commodities (such as taxi rides and spare rooms) to distributed ("platformised") business models where the perceived value grows together with and thanks to the number of users that are using the platform.

The creation of network effects involves offering strong, identifiable value. In the case of the EOSC, the challenge is to leverage assets (data, services, …) to offer new opportunities for EOSC users and suppliers. Viewed as a "platform", the EOSC can provide the business processes to facilitate interactions among the various actors present in the ecosystem (e.g. service and data providers and users). The PDT process supported exploration of ways in which this could be achieved, in scenarios such as those described by the ELIXIR and ASTRON use cases. The overall context for applying the platform design process to the EOSC was defined prior to beginning the PDT work and is stated in the Introduction to this paper.

In the case of the EOSC, the platform design process is focussed on capturing additional value in an *existing* ecosystem by improving understanding of the identities, motivations and contexts of its key participating entities. The EOSC ecosystem does not operate as a free-market environment but rather one in which public policy is a very significant driver and constraint.

A detailed description of the outputs from each step of the platform design methodology is included in Appendix III. The PDT process can be summarised as follows:

- identifying key roles in the EOSC "ecosystem" and the gains they aim to obtain

[36] The Platform Design Toolkit is produced by Boundaryless, who also provide consultancy and mentoring support

[37] Business Model Canvas and STOF business model framework

- mapping key relationships between the roles and analysing them with the aim of identifying those interactions with the greatest potential to derive value from the platform: improvements the platform strategy could bring to each transaction to make it *easier, cheaper and faster*
- identifying the support and enabling services which would allow entities to improve performance of their current role or, over time, to evolve and embrace new opportunities. This is where opportunities lie for the EOSC to add value
- assembling the roles, the transactions and the learning opportunities into an experience that synthesises the core value proposition essential for the platform strategy.

In modelling the ELIXIR and ASTRON use cases using the methodology, the key roles identified were Research Service providers (including publicly funded Research Infrastructures, pan-European and national e-Infrastructures), Research Support Providers (including support staff at Research Infrastructures, e-Infrastructures and universities (e.g. data steward, computing engineer), and Research Asset Consumers (including researchers (as consumers of existing research outputs), research performing organisations, research communities and communities of practice). The modelling identified the following improvements which could be made to aspects of transactions compared to the current situation:

- provide templates for use case description, to better extract use cases from the demand side to the supply side
- offer one-to-many interaction with providers (i.e. advertise demand), to allow a broader and more standardised response from suppliers to demanding parties
- standardise metadata, to enable a more standardised discovery of data, software and services and allow for easier comparison
- standardise access terms, to create a more consistent and predictable access and provision of services
- provide standard contracts, to bundle best practices and increase professionalism in the public sector
- provide information on funding solutions and conditions, to allow for better signposting and lower the threshold for low-capital research groups
- ease cross-border use of research assets such as storage and compute services, to support international organisations' research
- provide quality and usage statistics
- facilitate aggregated usage across wider sets of researchers
- support networks of research support providers for knowledge exchange
- increase quality and availability of research support
- provide user guides and best practice templates
- simplify services – e.g. intuitive interfaces, to lower the threshold for emerging research disciplines and lower the costs of training and support
- support interoperability of services for composability
- support rating of research assets and artefacts – user feedback.

Several of these points are already recognised as gains the EOSC should deliver and are being, or will be, addressed by EOSC-related projects, but the platform design approach helped to highlight that an increased volume of transactions (i.e. use of the EOSC for matching user needs and supplier offers), and greater and more efficient use of publicly funded research assets can be expected to result from making these improvements. The work clearly identified the importance of the role of Research Support Providers and the value they could add to the EOSC.

Opportunities identified for entities to improve and develop include:

For Research Support Providers

- Challenge at platform onboarding stage: finding suitable providers for the scientific cases they want to support, quantifying needs for these cases
    - Platform potential to address: consultancy to assist in articulating demand, catalogue of services, publication of scientific challenges (demand-side advertising)
- Challenge getting better in the current role: standardise framework agreements
    - Platform potential to address: develop framework agreements (e.g. between Research Infrastructures and e-Infrastructures to allow flexible/scalable use)
- Catching new opportunities: provide support outside traditional user community
    - Platform potential to support/encourage: promotion channel and matchmaking platform to find communities that need support.

For Research Service Providers

- Challenges at platform onboarding stage: Listing services, pricing services, promoting services to a wider audience
    - Platform potential to address: provide cost/pricing models, service descriptions, guided onboarding and reputation statistics (quality, usage, user satisfaction)
- Challenges getting better in the current role: provisioning for-pay services, leveraging local funding schemes for provisioning
    - Platform potential to address: funding scheme knowledge base, updates to regulations, guidelines and best practices, service procurement support (i.e. expert advice and administrative support, including potentially relating to aggregated procurement)
- Catching new opportunities: maintaining in-house status whilst being able to respond to new opportunities
    - Platform potential to support/encourage: provide framework that allows preserving in-house status on a wider scale.

For Research Asset Consumers

- Key challenges at platform onboarding stage: finding required assets, estimating trustworthiness of assets

    - Platform potential to address: catalogue of services, feedback and community support
- Challenges getting better in the current role: integrating solutions, receiving invoices
    - Platform potential to address: consultancy solution architecting, brokering to (local) funding schemes
- Catching new opportunities: becoming a data provider, establishing collaborations to improve existing software or develop new applications, stepping up as an expert supporter
    - Platform potential to support/encourage: supporting a human network, data lineage and ingest tools, dynamic results from EOSC observatory.

It can be seen that the EOSC has the potential to support each of these key roles to evolve in ways which allow them to offer more value, or to derive more value from the EOSC. Overall, the EOSC can support better science and more effective use of research assets.

4.2 Discussion

In terms of the overall "platform experience", the EOSC currently supports providers to onboard their services, while users can browse the catalogue and request them either directly via the portal or indirectly via the providers' websites. This type of experience works well when the user has a clear understanding of what they want to achieve and is able to identify the needed service. In the scientific computing sector, however, it is common for the user to need guidance and support to make optimal use of the available resources and technology. This need was highlighted by the platform design analysis, which modelled research organisations (Research Infrastructures and e-Infrastructures) discussing complex and large-scale needs, showing the role of Research Support Provider to be one of the primary beneficiaries of the EOSC portal. While this type of interaction is supported by the EOSC-hub project for specific cases (e.g. see Thematic Services and Competence Centre), it is not yet a capability offered by the EOSC Portal. The modelled "platform experience" also highlighted the important step of identifying a funding source to pay for the service usage, and of understanding the incentives for the research support providers to operate in the EOSC ecosystem.

The EOSC platform modelling complemented the use cases described in the previous chapter by abstracting the related model of interaction and value exchange. The focus was on publicly funded Research Service Providers, Research Support Providers and Research Asset Consumers, although the platform design methodology could be used to model numerous other transactions or business processes in the EOSC. Other models of interactions such as the transactions surrounding Research Asset Producers and Commercial Service Providers, were in fact analysed during the work but are not reported here as they were not in scope.

The modelling showed that the EOSC as a platform could be developed to support the matching of user needs and supplier offers, helping to make more efficient use of existing research assets, but that it does not currently provide the full range of capabilities required to do so for complex, demanding use cases such as those illustrated by ELIXIR and ASTRON. In particular, the current EOSC portal focusses more on facilitating selection of services from a predefined list and does not support the negotiation required for complex solutions, although this capability could be developed. The recommendation from platformisation experts is to rely on entities in the ecosystem (e.g. Research

Service Providers, Research Asset Consumers) to "orienteer" the process of populating the EOSC as much as possible themselves, as this not only reduces costs for the platform but also lowers frictions and provides the opportunity to capture the "known unknown": the community themselves know the information they wish to provide or need to find.

The modelling activity also helped to clarify that Research Support Providers - situated in research institutions, Research Infrastructures and e-Infrastructures - are important for realising value from the EOSC. It is likely they will be amongst the primary users of the EOSC portal: they have expert capabilities and can consult other Research Support Providers[38] on the best solution for their specific challenges. In particular, when Research Asset Consumers wish to exploit large-scale data, they require both specialised consultancy and support, and the ready availability of cross-border research assets such as storage and compute services.

In the next chapter these insights are combined with those from the use cases and discussed further

[38] It should be noted that the same organisation can play multiple roles. Here, a research community (ELIXIR, in this case) acts as a Research Asset Consumer when approaching the EOSC as "demand organisation", and as a Research Support Provider when enabling research groups to exploit research assets; on the other side, an e-Infrastructure can act as Research Support Provider when providing consultancy, demonstrators, etc, while it acts as Research Service Provider when offering technical services from its in-house infrastructure.

5. Insights and Recommendations

5.1 Introduction

The use cases and platform design work were very different activities, but they produced complementary insights: the use cases revealed, as intended, the challenges and issues involved in matchmaking demand and supply in demanding cross-border service use cases; the platform approach supported analysis of the motivations and desires of entities likely to be involved in the demand and supply of research services, and highlighted possible opportunities for realising greater efficiency and benefits. The findings were corroborated by the session on Issues in Cross-Border Consumption of Resources in EOSC which took place (virtually) on 19 May 2020 as part of the EOSC-hub week[39], where it was observed that the main barriers to cross-border service provision are not technical but legal, financial, organisational and regulatory. They include access policies and required levels of assurance; VAT; GDPR and IPR restrictions; sustainability of services; remit and imposed restrictions on funding models and business models; SLAs and OLAs; and are generally quite intractable and time-consuming to address. It was also observed that it is challenging to create opportunities for redistributing funds to national providers because service providers are bound to their given national and/or community mandates.

EOSC-hub Deliverable D12.2[40] studied procurement demand scenarios using four case studies examining, respectively, the use of vouchers, virtual access, framework procurement and public-to-public cooperation with cost recovery mechanisms. The last of these examined the possible application of the exemption to the EU public procurement directive 2014/24/EU[41] for public-to-public cooperation in the context of the ELIXIR use case. The reader should refer to D12.2 for full discussion of this case study, but it concluded that research institutions collaborating to provide services may be able to meet the conditions to satisfy the public-to-public cooperation exemption, meaning they could avoid conducting a competitive tender process. An important condition for this exemption was that the cooperation should be intended to achieve common objectives, and not just be a simple reimbursement of costs. Other types of exemptions are mentioned in Section 4.3.1.2 of D12.2, e.g.: sole-source provider, lack of competition, extreme urgency circumstances, contracts awarded on the basis of an exclusive right, low-value contracts.

In the following sections, these conclusions from the Cross-Border Services session and from D12.2 are combined with the work described in the previous two chapters to produce insights, and some recommendations are proposed for addressing them.

5.2 Rich Matchmaking Capability

The use cases highlighted that for demanding or high-volume cases, complex information needs to be accessed and exchanged before transactions (i.e. contracts for the supply of publicly funded compute and storage resources) can be concluded. Non-community-specific EOSC services such as those for

[39] https://eosc-hub.eu/events/eosc-hub-week-2020-goes-virtual
[40] https://www.eosc-hub.eu/news/participate-eosc-hub-procurement-and-business-models-consultation
[41] https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A32014L0024

federated AAI, metadata handling and PIDs are important, but the portal is required not so much for increasing exposure within one's own domain as for reaching out to other domains. Human networking with other domains was also recognised as important.

The majority of research in Europe is funded nationally. It is challenging to coordinate it to meet the needs of international research collaboration supported by depletable capacity such as storage and compute, particularly for multidisciplinary research collaborations. Funding sources are varied, complex and involve a large number of different rules, which contributes to suboptimal use of member states' investment in research services. Research use cases could be aided by providing "choreography" of national and European funding schemes and eligibility criteria across borders. A knowledge base of national funding opportunities could be created for international organised communities, signposting potentially suitable routes to secure viable funding, helping to reveal more of the research funding map. From the point of view of the EOSC as a platform, this opens the possibility of participation by agencies or professionals which specialise in providing assistance for dealing with funding rules and processes.

The EOSC can help to provide the necessary framework to facilitate matchmaking of use cases with offers from different providers but to do so the technical and commercial information it provides needs to go beyond what is currently being implemented. To develop as a federation of research infrastructures and e-Infrastructures, the EOSC needs to standardise and optimise processes, regulations and agreements involved in transactions, although the complex information involved also needs not to be oversimplified. Important suggestions which emerged from the use cases and PDT work include the ability to clearly distinguish types of value-added services, by providing services listings in more numerous, standardised and granular categories which permit different forms of value-added support to be distinguished (e.g. commodity services; different types of added-value services such as consultancy, collaboration, commodity resources including support; etc) as a first step towards reducing complexity; providing demand-side advertising of needs and challenges (such as those illustrated by the ELIXIR and ASTRON use cases) to provide the means for entities or research groups to attract attention from potential providers of solutions to their needs; and supporting the use of supplier reputation, as a driver of quality[42]. Viewed as a platform, the "population" of information in the EOSC should be driven and provided by EOSC users and service providers based on their needs, reducing the cost of the EOSC but also ensuring that the information it provides meets the needs of its users.

Rapid transactions are a key to success of the EOSC. The EOSC needs to put demand and supply in contact efficiently and consistently by acting as a platform to standardise and structure dialogue and optimise transactions: matchmaking too often currently relies on chance and needs instead to become a more standard process drawing on up to date, accurate and comprehensive information and with a realistic prospect of success. The functions and activities involved would not necessarily all need to be visible to end users but would nonetheless be providing significant value to support research collaboration. The PDT process helped to identify several support and enabling services the EOSC could provide to address this need in addition to those already mentioned above (some of which are touched on further below), including developing framework agreements; providing pricing models,

[42] For example, along the lines of tripadvisor.com

guided onboarding and reputation statistics (quality, usage, user satisfaction); updates to regulations, guidelines and best practices; procurement support and providing a framework which allows preserving in-house status on a wider scale; consultancy solution architecting; brokering to (local) funding schemes; and supporting human networks. Barriers and complexity relating to regulation (such as data protection, licencing and procurement) cannot necessarily be fully removed by the EOSC but it has a role to play in reducing them through enabling services and activities such as information services and tools (e.g. dashboards), consultancy support and advocacy.

Insight 1:

Rapid transactions are a key to the success of the EOSC. Capabilities are required of the EOSC as a platform, to support negotiation and matchmaking for complex and demanding use cases, including:

- standardised and optimised processes, regulations and agreements involved in transactions matching supply and demand
- services listings in more numerous, standardised and granular categories which permit different forms of value-added support to be distinguished (e.g. commodity services; added-value services such as consultancy, collaboration, commodity resources including support; etc)
- demand-side advertising of needs and challenges
- support for the use of supplier reputation, as a driver of quality
- a knowledge base to signpost information about national funding sources.

Recommendation 1:

The EOSC has the opportunity to make matchmaking a standard process with a high rate of success. It should provide the capability to put demand and supply in contact efficiently and consistently, including for complex use cases, and facilitate negotiation, by acting as a platform to signpost information, standardise and support dialogue and optimise transactions.

5.3 Research Support Providers as Primary Users

Information, or the extended range of functions recommended in the previous section, is a necessary but not sufficient part of unlocking the potential of the EOSC as a platform which orchestrates matchmaking between researchers and services. The use cases and the PDT work both showed that the role of Research Support Providers throughout the process of matching supply and demand is key to making use of information, to easing the effort and difficulty involved and to supporting the long tail of users to collaborate more strongly in research communities: the complexity and individual nature of many of the transactions which need to be encouraged mean that the EOSC as an information source, however rich, is not sufficient to deliver the “mass personalisation” potential of the platform. The knowledge and expertise of qualified, experienced Research Support Providers is needed to successfully match demand to supply, reducing friction in transactions. They can offer specialised support backed by specialised infrastructure to help users to understand the implications of different choices, build demonstrators, assess technologies, and help arrange use of services whilst

protecting users from lock-in and supporting movement between providers. Research Support Providers have an established place in the publicly funded e-Infrastructures, highlighting the e-Infrastructures' role as providers of expert technical support and consultancy. The contribution of Research Support Providers is also potentially important to show that a public-to-public cooperation aims to achieve common objectives, rather than simply reimbursement of costs, which is one of the conditions that could qualify for exemption from public procurement rules[43].

Research Support Providers would be amongst the most intensive users of the EOSC, challenged with becoming familiar with its offers or demands (as appropriate), and matching them up in the most effective way. They could be considered as proxies for Research Asset Consumers, "solution architects" who speak on behalf of those consuming the value of the EOSC. The capabilities of the EOSC discussed in section 5.2 would help to support them in their work.

The use cases and the platform design work described in this Briefing Paper illustrate the importance of the role of Research Support Providers in the process of bringing users and providers together successfully to allow more and better science to be performed. They thus form part of the value proposition of the EOSC. The creation of the EOSC could stimulate more entities supplying Research Support Providers to board the EOSC and provide their services to users, and certification based on skills and results could be developed to provide quality assurance and professional recognition. The creation of an EOSC Coordination Point providing a federated European network of distributed expert teams is recommended. This would be along similar lines to the OpenAIRE NOADs[44] or the EGI NILs[45], mobilising existing Research Support Provider resources, using resources from national e-Infrastructures (funded nationally) as a basis. A small EOSC coordinating team for this network could be funded by the EOSC, for example as part of the EOSC Secretariat.

Insight 2:

Qualified, experienced Research Support Providers – providers of expert technical support and consultancy - are key to reducing friction in research transactions. Their input is required to successfully boost research collaboration and make more effective use of research assets, for example to

- help users to understand the implications of different choices
- build demonstrators and assess technologies
- help arrange use of services whilst protecting users from lock-in and supporting movement between providers.

Recommendation 2:

Create an EOSC Coordination Point providing a federated European network of distributed expert teams, mobilising existing Research Support Provider resources, using resources from national e-

[43] See D12.2 for a more detailed discussion of the conditions of the public-to-public cooperation exemption
[44] https://www.openaire.eu/what-is-the-openaire-network-noads
[45] https://wiki.egi.eu/wiki/NGI_International_Liaison

Infrastructures (funded nationally) as a basis, and provide a small EOSC coordinating team funded by the EOSC, for example as part of the EOSC Secretariat.

5.4 Shared Resources

The ELIXIR and ASTRON use cases illustrate that there is a need for cross-border use of rivalrous resources such as storage and compute services (Shared Resources[46]), as was previously identified during the formulation of the EOSC-hub proposals for the EOSC Federating Core (see Appendix I). The capability to organise the exploitation of large-scale distributed datasets requires compute and storage resources to be adequately provisioned and integrated. Funding their cross-border use is complex and challenging, but to support international research collaboration a means needs to be found of doing this in a reliable manner which involves a proportionate amount of effort for research communities. The use cases showed that existing mechanisms such as Virtual Access, procurement frameworks and vouchers, whilst they each have their merits, do not necessarily provide suitable solutions. They also demonstrated the caution of publicly funded service providers with respect to the public procurement rules, causing suppliers to mostly provide services only to organisations with in-house status.

There needs to be a cost recovery mechanism whereby the research infrastructure (or other public organisation requiring services) reimburses the costs of one or more of the publicly funded national e-Infrastructures in return for the provision of digital services, although the e-Infrastructures' systems and resources for cross-border billing and contracting also need to be fit to meet requirements or perhaps alternatively simplified by standard contracts or central settlement for services. Analysis in D12.2 shows that a public-to-public cooperation exemption from public procurement rules could be a viable option. The EOSC should investigate this more and provide guidelines and framework structures to facilitate cross-border public-public cooperation. Supplementary national budgets to co-fund use of (own or other countries') national infrastructures may also be helpful, to provide additional demand-side funding to complement, or top up, the supply-side funding nationally, assuming that the proper mechanisms are in place to use them in public-to-public service provision. If member states' contribution to the EOSC Partnership were formed partly of in-kind donation of resources by member states to the EOSC, it would create a tighter coupling between national e-Infrastructures and the EOSC and provide a catalyst for cross-border use of services. There may be a role for the use of tokens by organisations which wish to consume digital services such as storage and compute (such as ASTRON, in the context of the use case discussed in this paper), for distribution to service providers in return for services. These options should all be investigated in more detail to assess their viability and potential benefits. Studies could also be undertaken to gather further empirical evidence of the macroeconomic and societal benefits of cross-border research funding. In the long-term, it is possible to envisage the emergence of arrangements where a country provides its facilities to another country in exchange for sharing the resulting research outputs and related reputation value. Open Science incentives and metrics would further encourage and support this.

[46] Resources including scientific outputs (local copies of data; applications, software, pipelines etc.) and the storage and compute hosting platforms needed to deposit, share and process them

As observed above, it is challenging to coordinate research funding to meet the needs of international research collaborations for services supported by depletable capacity, such as storage and compute, particularly for multidisciplinary research collaborations. Shared Resources not only need to be sustained, but also a way needs to be found for international research communities to pay service providers which have national scopes and remits. A more frictionless marketplace needs to be created, in which funds flow to those service providers which deliver the most value. Current EC or international funding schemes are not well suited to funding cross-border, cross-disciplinary consumption of depletable resources such as storage and compute. There is at present no straightforward means for a research group to pay with EC research funding (for example European Research Council funding) for use of existing e-Infrastructure, but it is possible for such funding to be used to build or develop new infrastructure. This increases fragmentation of digital service provision and reduces utilisation of existing e-Infrastructure, highlighting the suboptimal use of funding. There needs to be a tighter coupling between research funding and e-Infrastructure funding, at the EU level and also on a national level.

Shared resources such as storage and compute are a fundamental part of the EOSC: as is the case with high-speed networking, there is a need to provide such services outside of the researcher's own domain. The EOSC Partnership needs to ensure appropriate coordination and synchronisation of member state and EC funding to sustain the whole EOSC ecosystem including the Shared Resources. The mechanics for funding the shared resources are complex. Centrally held European money is difficult to allocate but has the advantage of being somewhat blind to geography and some EC co-funding of access to national infrastructures may provide the necessary "glue". On the other hand, nationally held funds or in-kind contributions may be easier to administer and should be considered for inclusion as part of the minimal viable EOSC. Dialogues should take place within member states between ministries, research funders and research service providers[47] to allow for a cohesive input and contribution across the operational support of pan-European research, representation in the EOSC Partnership, and the broader EU research policy.

Pilot cross-border research collaborations could be supported to help assess the potential value of data and other resources for new use cases. The EOSC could also take some other steps to support greater efficiency: liaising with other research funding schemes (e.g. ERC) to avoid research groups using grants to build services which are already available from EOSC providers; and encouraging thematic communities to identify resources they could share with other communities. It is challenging for research disciplines to be able to express their long-term infrastructure needs so that they can be incorporated into infrastructure capacity planning and funding. It would help if project proposals were encouraged to consider all requirements for the work, including estimated infrastructure needs and how their costs will be met. Assessment of proposals could be extended to cover whether such considerations have been included. Some national roadmaps have begun to require this, but it would help to apply it more widely including in Horizon Europe.

[47] The e-IRG report "National Nodes - Getting Organised; How Far Are We?" recommends (page 22) "Members states and associated countries should continue to increase the level of coordination between and consolidation of the various national players on e-Infrastructure provisioning." See http://e-irg.eu/catalogue/eirg-1006.

Insight 3:

Shared Resources such as storage and compute are a fundamental part of the value-add of the EOSC. A means needs to be found to fund their cross-border use in support of international research collaboration. Measures which may help include:

- studies to obtain further empirical evidence of the macroeconomic and societal benefits of cross-border research funding
- support of cross-border research collaborations organised in pilots to help assess the potential value of data and other resources for new use cases
- developing guidelines and structures that facilitate cross-border public-to-public service collaboration with cost reimbursement (e.g. public-to-public cooperation exemption from public procurement rules)
- providing supplementary demand-side funding for use of national e-Infrastructures, including their cross-border use, to complement or top-up national supply-side funding
- payment of member state contributions to the EOSC Partnership partly in the form of in-kind donation of resources by member states to the EOSC, to create a tighter coupling between national e-Infrastructures and the EOSC.

Recommendation 3:

The EOSC Partnership needs to ensure appropriate coordination and synchronisation of different funding streams (e.g. from member states, EC, research performing organisations and projects etc) to sustain the whole EOSC ecosystem including the Shared Resources. To underpin this policy, dialogues should take place within member states between ministries, research funders and Research Service Providers to allow for a cohesive input and contribution across the operational support of pan-European research, representation in the EOSC Partnership, and the broader EU research policy.

6. Conclusions

The work described in this Briefing Paper set out to study the cross-border provision of publicly funded rivalrous resources such as compute and storage, motivated by the EOSC-hub proposals for the Federating Core which identified the need for such services to form part of what the EOSC offers and sustains.

The EC's stated objective for the EOSC is for it to be a fundamental enabler of Open Science and of the digital transformation of science, offering every European researcher the possibility to access and reuse all publicly funded research data in Europe, across disciplines and borders, leveraging past investment in research data infrastructures to add value in terms of scale, interdisciplinarity, and faster innovation[48].

The EOSC has the opportunity to reduce barriers and enable frictionless, easy access to data and related services across geographic boundaries and disciplines so research communities can better connect with suppliers, users and funders, and to promote a cross-border multidisciplinary environment where past, present and future investments in resources and infrastructures can be efficiently leveraged and benefit from economies of scale.

The work has demonstrated ways in which the EOSC can develop to meet these objectives but it has also highlighted the suboptimal use of public funding: long-term strategies on science and infrastructure are not necessarily formulated in a joined-up manner, and EC and member state funding do not interoperate efficiently. The majority of e-Infrastructure investment is still funded nationally. It is challenging to coordinate it to meet the needs of international research collaboration. On a thematic level coordinated activity does take place, but it is particularly challenging to coordinate funding for e-Infrastructure services in support of multidisciplinary research collaboration. The complexity and fragmentation of funding rules for cross-border research activities is a recognised issue highlighted by the EOSC, which needs to be addressed, but it has received limited attention during the implementation of the EOSC to date. The GÉANT IaaS Framework[49], the OCRE project[50], the EOSCpilot Policy Activity[51] and EOSC-hub's procurement deliverables D12.1 and D12.2 are the main activities which have provided solutions or studied relevant options. The insights described in chapter 5, and the conclusions in D12.2, point the way to possible options for funding and providing Shared Resources across borders, which should be investigated further.

The insights and recommendations were the subject of a public consultation from 14 September to 9 October 2020, supported by a webinar on 15 September. The proposals were also discussed in other fora including the EOSC Architecture Working Group on 16 September. The consultation survey received only two responses but the feedback received from all sources was positive overall and supportive of the proposals. One correspondent observed that the Shared Resources are needed in the EOSC *from the start,* because [access to] data on its own is not enough. One of the survey

[48] https://ec.europa.eu/research/openscience/pdf/swd_2018_83_f1_staff_working_paper_en.pdf
[49] https://clouds.geant.org/geant-cloud-catalogue/geant-cloud-catalogue-iaas/
[50] https://www.ocre-project.eu/
[51] See https://eoscpilot.eu/content/d36-final-policy-recommendations

responses pointed out that the EOSC Association and its membership fees are not intended to be used to run or pay for the EOSC core, and that discussions about the EOSC core and how to finance it need to take place between the EOSC Association, the European Commission and the EOSC Steering Board (i.e. the EU member states). Small updates were made to the recommendations in Chapter 5 reflecting this clarification, although with regard to the EOSC Secretariat, whatever governance and funding is finally agreed for it, a realistic view should be taken of the numerous activities it will need to perform and the importance of providing adequate funding for it to successfully support the functioning of the EOSC.

The EOSC would not fulfil its mission just by organising a web of FAIR data; the EOSC has clear potential to act as a platform stimulating further research collaboration, producing more and better science to the benefit of all, but it involves the practical implementation steps recommended in the previous chapter, in particular accepting the cost of sustaining shared resources whose value has been illustrated by the use cases. The EOSC user experience is key to the success of the EOSC: researchers want tailored consultancy services rather than an off-the-shelf service. A critical success factor of the EOSC is to optimise and choreograph the fulfilment process: *what* is supplied may be a commodity, but *how* it is supplied is not. Research Support Providers, particularly those playing an established role in the national and European e-Infrastructures, have a crucial role to play in this, acting as trusted advisors. The basis for the platform already exists within the research ecosystem and can be built upon. The e-Infrastructures are recognised as the horizontal foundation of the EOSC; the work described here confirms this view.

In a sense, implementing the EOSC switches from the European and national funding authorities providing solutions for the research community, to a policy which enables the community to self-organise. The platform approach shows that to achieve this, policy makers and funders need to give more consideration to potential and existing flows of value than to the political agenda. They need to facilitate the ecosystem to self-organise, derive greater value from relationships, and evolve according to its own needs, although the research community can still be influenced by policy makers towards desired outcomes and behaviours.

Subsidiarity – dealing with issues at the most local level possible – provides an incentive for member states and national EOSC stakeholders to co-execute and co-run the ecosystem-empowering services – including the shared resources - that help other entities in the ecosystem to onboard, improve and eventually discover transformative opportunities inside the ecosystem.

The demand exists for cross-border use of assets. Many of the necessary services and consultancy resources are also in place, although they operate mainly within national or community boundaries and the importance of their sustainability beyond the service development phase needs to be better recognised and supported. To allow more and better science to be performed, there is a clear incentive for governments and national funders to facilitate and sustain cross-border use of assets.

This Briefing Paper has begun to explore the EOSC as a platform but the platform toolkit and methodology could be deployed further to derive a more profound and complete understanding of the potential of the EOSC, including some of the recommendations suggested above for its development but also, for example, its use to advance Open Science since by design the focus in this paper has been on compute and storage. More use cases could be explored, perhaps to analyse the

ecosystem from the point of view of long-tail users or a repository provider or to study producing FAIR research outputs; there could be examination of other roles in the ecosystem, such as commercial service providers; and the PDT approach could also be used to further examine the place of training, quality assessment and trust in the EOSC. Use cases such as those explored in this paper could also be extended to consider more countries and national providers, to gather more information about national services and processes. Further study of the different types of reimbursement streams which could be arranged for public-to-public service provision would also be worthwhile.

Appendix I EOSC-hub Federating Core Proposals and Comparison with Tinman Proposals

I.1 EOSC-hub Federating Core Proposals

EOSC-hub Task 2.3 published initial proposals for the EOSC Federating Core in July 2019[52]. Following public consultation, updated proposals were published in November 2019[53]. These are summarised below.

Requirements of the EOSC

The EOSC should provide in a coordinated manner, across disciplines, the provisioning of capabilities that are generally applicable to data lifecycle management:

- discovery and reuse: provide a means for universal and versatile discovery and sharing of resources[54], through a portal, and with inclusive and transparent policies for access
- processing and analysis: provide for common user needs for generic storage and processing facilities for data management and analysis, such as high-performance and high-throughput distributed compute capabilities, for researchers to manipulate resources to which they have been afforded access via the EOSC
- data management, curation, preservation: complement what is provided by research institutions and communities, according to the subsidiarity principle, and provide researchers with a working space where they can use EOSC resources collaboratively. Research institutions, communities and infrastructures should remain the main custodians of research data, quality and FAIR policies
- access/deposition and sharing: widen access to data produced, curated and preserved by national and European research communities and enable the reuse of research communities' FAIR data and data analytics tools.

In the interests of delivering opportunities and efficiencies, the EOSC needs to interoperate with other infrastructure initiatives in Europe and other regions, including EuroHPC[55].

[52] https://www.eosc-hub.eu/sites/default/files/EOSC-hub%20Briefing%20Paper%20-%20EOSC%20Federating%20Core%20Governance%20and%20Sustainability%20v1.0_0.pdf

[53] https://www.eosc-hub.eu/sites/default/files/EOSC-hub%20Briefing%20Paper%20v2.0%20-%20EOSC%20Federating%20Core%20v0.3%20%28consultation%20comments%20and%20responses%29%20%282%29.pdf

[54] Defined in the EOSC Portal glossary, November 2019, as *any asset made available (by means of the EOSC system and according to the EOSC Rules of Participation) to EOSC System Users to perform a process useful to deliver value in the context of the EOSC. EOSC Resources include services, datasets, software, support, training, consultancy or any other asset.* See https://www.eosc-portal.eu/glossary

[55] https://eurohpc-ju.europa.eu/

EOSC Federating Core

The Federating Core required to support this vision of the EOSC is defined below.

The **Federating Core** is a fundamental asset of the EOSC, composed of the technical, human, policy and resource elements required to facilitate, monitor and regulate as appropriate day-to-day transactions across the federation.

The Federating Core should deliver three capabilities:

(1) **Hub Portfolio**: The activities and tools necessary to provide coordinated access to and management of resources[56] provided in the EOSC Shared Resources or the Service Portfolio. EOSC resources are expected to be delivered at national and European level, together with the support and expertise necessary to address complex digital needs of the EOSC user communities. The Hub portfolio delivers the EOSC "**federating tier**".

(2) **Compliance Framework**: The Rules of Participation, the Interoperability Framework, the Service Management System and other policies and processes for suppliers and users to engage with the EOSC. The Compliance Framework constitutes the EOSC "**regulatory tier**".

(3) Shared Resources: Resources including scientific outputs (local copies of data; applications, software, pipelines etc) and the storage and compute hosting platforms needed to deposit, share and process them. The shared resources realise the EOSC "**resource tier**".

The Federating Core is complemented by the **EOSC Service Portfolio** which provides additional added-value services which exploit the Federating Core, are delivered by providers external to the EOSC according to independent provider-specific business models, and are discoverable through the EOSC Portal.

The composition of the Federating Core and the EOSC Service Portfolio will be driven by EOSC-defined Rules of Participation[57], technical and policy requirements that will define the EOSC conformance requirements for providers.

The set of capabilities delivered by the Federating Core is defined by the EOSC governance, and the costs of its delivery shall be sustained by EOSC funding.

Initial proposals of the functions and resources which might comprise the three elements of the Federating Core are provided below[58].

[56] Defined in the EOSC Portal glossary as *any asset made available (by means of the EOSC system and according to the EOSC Rules of Participation) to EOSC System Users to perform a process useful to deliver value in the context of the EOSC. EOSC Resources include services, datasets, software, support, training, consultancy or any other asset.* See https://www.eosc-portal.eu/glossary

[57] Defined in the EOSC Portal glossary, November 2019, as *the principles defined by the EOSC Governance to drive the processes enacting an actor to play the role of EOSC System User (and any specialization of it)*.

[58] Descriptions of most of these can be found in the initial EOSC-hub Briefing Paper on the Federating Core, July 2019 (section 3.1) with information on more recent additions in the November 2019 Briefing paper (section 3.3.2).

Hub Portfolio

The list of functions proposed to constitute the Hub Portfolio includes

- EOSC Portal
- EOSC Support Services, including training, competence centres and knowledge bases
- EOSC AAI
- EOSC Data Transfer services
- EOSC Monitoring
- EOSC Accounting
- EOSC Configuration Management Database (CMDB)
- Collaboration Software
- Operations Portal
- EOSC Security policies and security coordination functions.

Compliance Framework

The Compliance Framework is proposed to comprise

- EOSC Rules of Participation
- EOSC Service Portfolio Management Tool
- EOSC Interoperability Framework
- EOSC Service Management System.

Shared Resources

The resources required from the EOSC Shared Resources differ from one discipline or community to another but possible capacities they could comprise include

- High bandwidth networking connectivity for high-performance access to EOSC data hosting nodes that provide storage and compute resources
- AAI services
- High-performance European distributed (federated) cloud storage environments for secure access, staging, downloading and deposition of large volumes of data across national, institutional and Research Infrastructure boundaries
- High-performance and high-throughput distributed (federated) compute capabilities for big data processing and analysis, including simulations
- A powerful search machine to support findability of scientific resources including data, tools, software and publications across many domains, and easily browsable federated dataset catalogues
- A repository of tools, services, software and workflows for data exploitation: simulation, analysis, enrichment and comparison of data from different national domains
- A catalogue of training materials and competence reference materials
- Open Science policy and practice recommendations for institutions and other EOSC stakeholders
- A code repository
- PID services
- Personalised workspaces for researchers, based on federated AAI.

Federating Core Value and Cost Model

Together, the Federating Core and the resources and research outputs provided with the coordination of e-Infrastructures and ESFRI projects and landmarks, would constitute a rich ecosystem which represents a significant part of the Minimum Viable Ecosystem (MVE) proposed by the second EOSC High Level Expert Group[59], to be complemented by resources provided to the EOSC by national research infrastructures. This ecosystem has high potential value for users and other stakeholder groups.

The EOSC needs to sustain the costs of providing the benefits of open data policies to a wider community of users. It needs to create the financial vehicle to cover the costs of the Hub Portfolio and Compliance Framework, and of provision and consumption of the Shared Resources beyond their originating communities. Coordinated provisioning and funding of the Federating Core is expected to bring economies of scale by aligning investments from member states with the compensation of marginal costs associated with cross-border usage of depletable resources and services.

I.2 Comparison with Tinman Proposals

The EOSC-hub Federating Core proposals were shared with the Sustainability Working Group and provided input into their Tinman document published in December 2019. The Tinman envisages the Minimum Viable EOSC including the EOSC-Core and the EOSC-Exchange, working with the FAIR datasets to be federated via EOSC, as shown in figure 3.

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

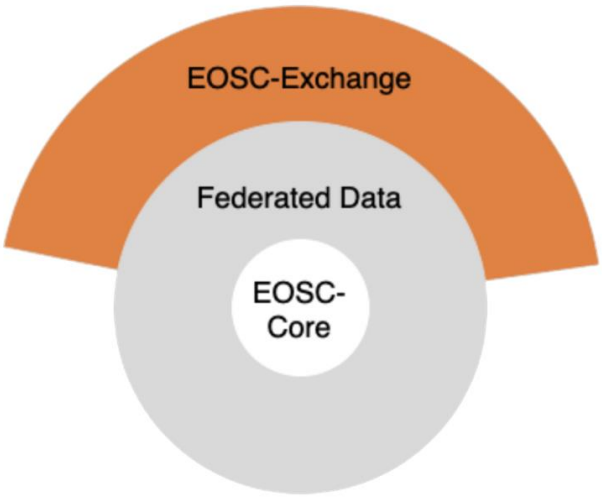

Figure 3: Schematic representation of key elements of the Minimum Viable EOSC[60]

According to the EOSC Executive Workplan 2020[61]:

- the EOSC-Core includes a restricted set of basic integration services. It provides the means to discover, share, access and re-use data and services but will not store, transport or process data, at least initially
- the EOSC-Exchange builds on the EOSC-Core. It is the place where rivalrous services (for example to transport, store and compute research data) will be made available.

A comparison between the MVE presented in the Tinman document and the EOSC Federating Core as proposed by EOSC-hub is provided in figure 4.

[60] Source: Sustainability Working Group Tinman document

[61] https://op.europa.eu/en/web/eu-law-and-publications/publication-detail/-/publication/ae215698-af7b-11ea-bb7a-01aa75ed71a1 See page 6

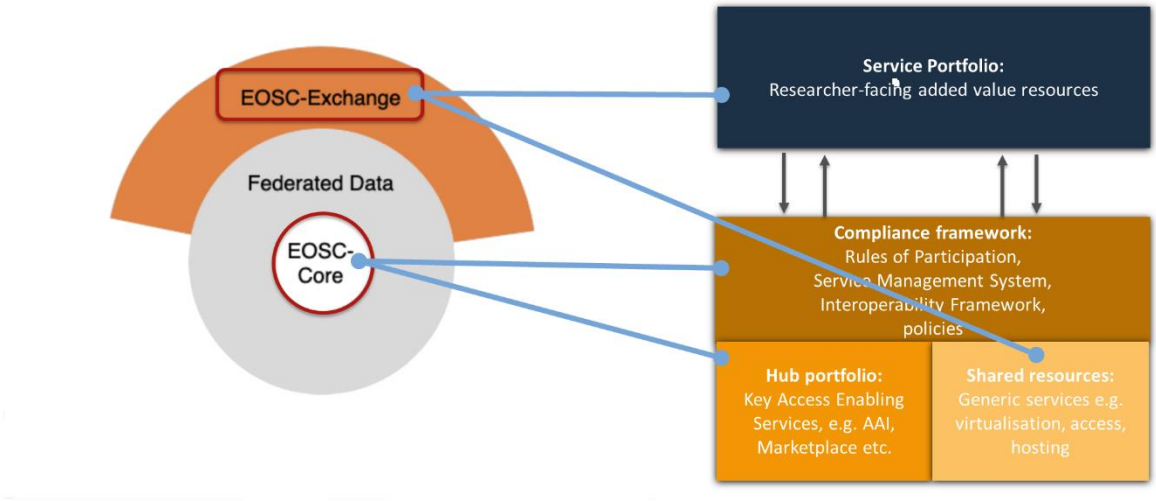

Figure 4: Comparison between Tinman Minimum Viable EOSC and EOSC-hub Federating Core

The comparison shows that the Shared Resources as proposed by EOSC-hub form part of the EOSC-Exchange proposed by the Sustainability WG Tinman.

The Sustainability Working Group is currently drafting an "Iron Lady" document, building on the Tinman document, which will provide an important contribution towards the definition of the Minimal Viable EOSC to be presented by the EOSC Executive in December 2020.

The EOSC-hub Federating Core proposals were also provided as input to the Architecture Working Group, forming a basis for its work (still ongoing) to identify the architecture components required for the EOSC.

Appendix II Use Cases

II.1 Use Case 1 – ELIXIR

Use case 1 concerns the ELIXIR community which is represented for the goal of this simulation exercise by the European Bioinformatics Institute (EMBL-EBI)[62].

By "ELIXIR community" here, we mean hundreds of research groups across Europe, each one usually formed of 5-10 people and led by a Principal Investigator (PI). Their typical research analysis project lasts between 3 and 6 months, but large collaborative analysis could take many years.

The simulation focussed on the science case of scale-out analysis: outsourcing the computing platform for hundreds of research groups to analyse ELIXIR datasets (order of Petabytes) in combination with their own data.

The analysis foresees the combination of public datasets with the group's own research data (private, with personal data for which they have obtained authorisation). Different research projects may access overlapping parts of ELIXIR public datasets.

Data storage or caching are often required from the start of a project; the computing may ramp up during the project depending on the maturity of the solution.

The need of computing for each analysis is the burst of computing from a couple of days to weeks depending on the degree of parallelism.

The simulation exercise involved four service providers located in different countries:

- CESNET (Czech Republic)[63]
- CSC (Finland)[64]
- INFN (Italy)[65]
- SURF (Netherlands)[66].

II.1.1 Description

The Scale-out Analysis science case is related to small research groups wishing to undertake a specific data analysis combining large public datasets with their own data. Typically, this will require:

- data to be transferred for analysis
- the analysis workflow to be able to scale out to be able to consume a large concurrent number of resources
- verification of the software stack, including the analysis workflow
- efficient access to large-scale public datasets and combination with private datasets from individual groups of researchers

[62] https://www.ebi.ac.uk
[63] https://www.cesnet.cz/cesnet/?lang=en
[64] https://www.csc.fi/en/csc
[65] https://home.infn.it/en
[66] https://www.surf.nl/en

- processed data or results to be transferred from the processing environment back to EMBL-EBI.

This work could last a few months and require thousands of vCPUs or hundreds of GPUs. Typically, the raw input data are of the order of 1 PB. Output data following analysis for movement back to the institute are of the order of 10 TB.

Institutional data will be exposed through solutions such as OneData, Globus Transfer, GridFTP to either expose the data for mounting into the cloud provider or to copy the data temporarily into the cloud provider.

The application is generally a high throughput computing (HTC) analysis as opposed to a supercomputing (MPI[67]) use case. Most current use cases use LSF[68] to manage access to the compute resources, but in the process of porting to cloud-based resources workloads are being transferred into Kubernetes-based models.

Various requirements were discussed during the technical session, specifically:

- the services to be operated
- the amount of resources (computing and data)
- authentication and authorisation
- the privacy and security of the data to be handled
- software stacks
- levels of support
- compliance to standards.

During the subsequent session, the commercial requirements were discussed, specifically:

- funding models
- legal and contractual requirements
- the challenges for the EOSC.

Through the discussions in the technical and commercial sessions, the service providers elaborated their offer to EMBL-EBI[69]. Two additional sessions to wrap-up and elaborate on the outcome of the exercise were held in the following days. In the next section the key elements of the outcome of this exercise are highlighted.

II.1.2 Technical Offerings

The service providers involved in the simulation exercise described what they were able and willing to offer to address the needs of the Elixir demand. In this section a summary of each offering is reported.

CESNET	**SURF**

[67] Message passing interface
[68] Load sharing facility
[69] Details are contained in https://docs.google.com/document/d/10Qx70EdgvYHfz9aNQYBc_W9k2HQd6VqiMCG-HYsCQBA/edit?usp=sharing

- Computing could be provided either on their MetaCentrum cloud, an IaaS infrastructure based on OpenStack, or on an HPC cluster - Storage: CEPH used in OpenStack - Data Transfer: gridftp or FTS - AAI: Support for EGI Check-In, ELIXIR AAI, national AAI - Support: best effort	- service would be provided through Spider, a fully managed HTC platform (SLURM Slurm scheduler), built on OpenStack, for I/O intensive processing with a focus on (sustained) embarrassingly parallel processing. The generic Spider platform (doc.spider.surfsara.nl) can serve the requirements as put forward by the use case (this is the nuclear setup). In addition this generic platform can be extended with external generic SURF services based on e.g., dCache scalable storage (disk and tape) to decrease the required staging area for active data and to manage costs (based on the required active data space, three setups were proposed: mounted space, green space and fully transient space). In addition it is possible to create custom Spider instances for this use case, e.g., to support private networks and Elixir IdP's according to the needs of the different user groups. The Spider platform could be extended with GPU nodes - Shared storage based on CEPH (accessed via CephFS) - Interoperable with scalable external storage systems (tape and disk) e.g., dCache, SWIFT and Data Archive over EVPN network (1200 Gbit/s) - AAI: Can supports AAI (via customization) - Support: Standard Helpdesk service (working days 9-17.00) with generic services; additional expertise hours can be purchased
CSC	**INFN**
Three 'service packages' were identified: 1. Non-sensitive data - Computing capacity on available on cloud (cPouta), cluster (Puhti) and HPC systems (Mahti and upcoming LUMI EuroHPC system) - Shared file system (Lustre) on cluster and HPC systems - Data can be stored in EUDAT Services (B2Safe) and bring into the computing through Eudat data transfers; or it can be outside EUDAT Services 2. Sensitive data - Computing at CSC secure cloud ePouta Data storage on CSC Sensitive data platform (in beta testing) - Currently capacity for computing may be limited - less than 1,000 nodes; maybe around 100 nodes	- IaaS Based on OpenStack, both computing and storage (block and object) - Planning to add docker support directly on bare metal (no VM) - Standard HPC environment also available - PaaS Based on INDIGO Orchestrators - AAI: ELIXIR AAI already supported, other OpenID solutions available - Support: Mon-Fri working hours, best effort during weekends

Requires setting up a special data transfer connection (multi domain VPN) 3. ELIXIR federated EGA solution - ELIXIR Finland node hosted at CSC working on setting up a node of the EGA infrastructure - Enables moving data stored at EGA to CSC with data custodian permission, and to computing at ePouta secure cloud - Integrated with ELIXIR compute platform, e.g. ELIXIR AAI CSC IT platforms have ISO 27001 certification	

II.1.3 Commercial Conditions

The four service providers highlighted the possible options they usually rely on and those suitable for the ELIXIR use case. These are summarised below.

CESNET	**SURF**
- Three different options available: 1. ELIXIR purchases the service directly from CESNET; 2. ELIXIR or EBI purchases the service through tendering process; 3. ELIXIR/EBI form a consortium and seek for EU funding - CESNET would not enter into an SLA which included very heavy penalties - Pricing heavily depends on the type of contract/project	- Typically three options are available: 1. Procurement of services against payment 2. Apply under the national funding scheme 3. Third-party funding under active EC-funded project - As the resource requirements, are assumed to be above the public procurement threshold with a policy that SURF does not respond to tenders, procurement by ELIXIR/EBI also presents an alternative cost by eating into SURF's 15% limit for revenues earned from non-members, they do not see any viable models where the ELIXIR/EBI legal entity can directly procure with SURF or apply under the national funding scheme - Third party funding under EC-funded projects give insufficient flexibility to fit the needs of the ELIXIR use case; - Any local Dutch ELIXIR group is however eligible for the national funding scheme or for procurement through their host institute, provided this is a SURF member organisation
CSC	**INFN**
- three different options available: 1. ELIXIR purchases the service directly from CSC (the Finnish ELIXIR node) as ELIXIR research collaboration	- two different models: 1. Pay-per-use model supported by a scientific collaboration 2. National funds

2. ELIXIR or EBI purchases the service through tendering process 3. ELIXIR / EBI form a consortium and seek for EU funding - Implementation of the case requires adjusting the service and for this reason CSC would propose a long-term contract	- possibility to establish a research agreement with the provider, no need for a public tender

II.2 Use Case 2 – ASTRON

This use case concerns the Radio Astronomy Competence Centre use case. ASTRON[70], the Netherlands Institute for Radio Astronomy, plays a similar role to that of EMBL-EBI for the Elixir community in Use Case 1.

Among other activities, ASTRON is responsible for the Low Frequency Array (LOFAR) telescope operations. It aims to offer a data processing service to the LOFAR community to help the generation of science-ready data products from data stored in the 'instrument' data archive. The latter is a distributed data archive with federated sites in Amsterdam (Netherlands), Jülich (Germany), and Poznan (Poland), but more data centres might be involved in the future, depending on the interest of countries joining the International LOFAR Telescope.

The simulation exercise involved three service providers located in different countries:

- CSC (Finland)
- INFN (Italy)
- SURF (Netherlands).

It is worth highlighting that one of the service providers (SURF) is already supporting ASTRON/LOFAR while the other two are in countries not participating in the LOFAR initiative for the time being.

II.2.1 Description

Given the volume of the archive (over 50 PB) and the typical data access pattern (i.e. data is accessed a few times with long periods in between), data is stored on nearline medium (tape).

Typically, datasets that are processed further are a few TB in size and single data products are tens to hundreds of GB in size. The processing has a low compute cycle per data volume ratio and lends itself well for High Throughput processing. At the moment, the main applications require POSIX access to data and do not benefit from accelerators although it is expected that the most compute-intensive task will soon (in a year or so from now) be ported to GPU which will significantly increase the efficiency.

Output data is still large although the highest-level data products, which are expected to be most in demand, will be significantly smaller in size. It remains to be demonstrated how much storage capacity is needed for this but the working assumption is that an online storage system of roughly 1 PB will be needed to provide immediate access to higher level data products (as generated by the service),

[70] https://www.astron.nl

possibly backed by a scalable nearline storage system. Long-term storage of the larger data volumes is in any case not expected to require very low latency and/or random access.

The operational model is still under discussion and will likely evolve over time. The initial phase is envisioned to be a managed service where:

- users apply for compute resources
- standard processing pipelines for generating higher level data products from data in the instrument archive are run on behalf of the user.

Later phases may include offering compute as a platform (closely integrated with the instrument archive) where users are allowed to submit their own workflows, possibly using pre-packaged images containing data analysis software packages.

ASTRON foresees three different activities, reported below according to their priority:

1. Running standard processing pipelines on behalf of the user
2. Science data repository for science level data
3. Compute as a platform.

The simulation exercise for this use case was shorter than the ELIXIR one due to time constraints, however most information was collected in advance and during both technical and commercial sessions all the relevant topics were discussed in depth. The most relevant technical items discussed were:

- the needs in terms of computing, storage and related services including operating systems, batch systems, virtualisation, containers, metadata
- authentication and authorisation
- support
- compliance to standards.

The commercial part focused mainly on the issue of funding, specifically on how to obtain funds either at European or national level.

II.2.2 Technical Offers

The service providers involved in the simulation exercise described what they are able and willing to offer to address the needs of the ASTRON demand. In this section a summary of each offering is reported[71].

SURF	**CSC**	**INFN**
- Service would be provided through Spider, a fully managed HTC platform (SLURM Slurm scheduler), built on OpenStack, for I/O intensive processing with a focus on	- Operate suitable services based on OpenStack with CEPH storage - GPU resources are also available	- IaaS Based on OpenStack, both computing and storage (block and object) - Planning to add docker support directly on bare metal (no VM)

[71] Details are contained in https://docs.google.com/document/d/1smGoW7lRsPQ3dHCunV_rMJ5yYaoNps4V7lp7eRFNNxc/edit?usp=sharing

(sustained) embarrassingly parallel processing - The same Spider options described for the ELIXIR use case are available for ASTRON - GPU nodes with NVIDIA V100 (32 GB) are possible Global file system CephFS optimised for high data availability, making it more stable for I/O intensive workflows and external dCache storage can be globally accessed and/or mounted - Access to a customisable data repository (SURF Data Repository) that has community support and support dedicated metadata schemes - AAI: Can supports AAI (via customisation) - Support: Standard Helpdesk service (working days 9-17.00) with generic services; additional consultancy hours can be purchased	- No shared file system is needed, so storage volumes that are connected to virtual nodes would be used - Data can be stored in EUDAT Services (B2Safe) and brought into the computing through Eudat data transfers; however it can also be outside the EUDAT services	HTC/HPC based solution with - HTCondor as a batch system and GPFS-based parallel file system with GridFTP based WAN transport solution. Access through Ssh-keys login. HPC could be used with a whole-node approach if needed/useful, and provide GPU card NVIDIA K40 - PaaS Based on INDIGO Orchestrators - Additional option: Bare metal V100 enabled Mesos cluster for running very big jobs → this is the recommended option for the ASTRON use case - AAI: various OpenID solutions available - Support: Mon-Fri working hours, best effort during weekends

II.2.3 Commercial Conditions

The main difference with the ELIXIR use case is that ASTRON/LOFAR currently has no standing budget to fund new external infrastructure resources, so the demand is necessarily subject to the possibility of attracting new EC/national projects funds. ASTRON do not see a viable model where users need to pay. However, the three service providers highlighted the possible options they usually rely on and those suitable for the ASTRON use case. These are summarised below.

SURF	**CSC**	**INFN**
SURF has a long-term relationship with LOFAR. The current policy is that data storage which is not part of an active compute job covered by an active grant, needs to be paid for. Computing is funded via NWO: every two years, a new application needs to be submitted; there is an informal policy that no more than 10%	There is currently not a Finnish LOFAR community, which makes availability of national funding unlikely Thus CSC would prefer a contract between LOFAR and CSC about the service covering all production costs. Common funding proposal for EC could be an opportunity, and CSC	INFN is not a member of the LOFAR initiative, and this makes things a bit more complex than for ELIXIR. However the two models already mentioned are still valid: 1) Pay-per-use model supported by a scientific collaboration, 2) National funds. There is the possibility to establish a research

of the economic capacity of the infrastructure can be granted to an individual project. Potentially additional grants under different funding instruments (national or international) could be used to complement computing grants	could also involve additional EUDAT CDI partners	agreement with the provider, no need for a public tender.

Appendix III Platform Design Process

III.1 EOSC Platform Design Process

III.1.1 Key Roles In EOSC

The PDT process involves identification of key roles involved in the EOSC "ecosystem" and the identification of the gains they aim to obtain. These are described here together with some of their characteristics. It should be noted that the same organisation can play multiple roles.

Role: Research Service Provider: for example, (publicly funded) Research Infrastructure, pan-European e-Infrastructure, national e-Infrastructure, research performing organisation, library

- **Goals** may include serving current constituents, achieving efficient use of available resources, ensuring visibility to potential thematic users, sustainability, fulfilling KPI goals set by their funding provider(s)
- **Potential gains from EOSC** include
    - Optimise volume operations to create economies of scale; ready-made service catalogues and order process would reduce operational overhead; optimise sourcing strategy between resource in ownership and commercially sources resources
    - scaling up to a larger user base; easy access to general building blocks needed by many (e.g. AAI, open source software stacks etc); potential to create new collaborations
    - recognition of being efficient and effective (reputation improvement, proven value-add); move to higher value-add services; potential to participate in new innovative projects; benefit from cost recovery mechanisms (along the lines of Virtual Access).

Role: Commercial Service Provider: for example, cloud service provider

- **Goals** may include producing profits, identifying market expansion opportunities
- **Potential gains from EOSC** include
    - offering standard/packaged solutions
    - access to new markets/bigger market share; "soft vendor lock-in": offering proprietary enhancements to standard interfaces, controlling de facto standards
    - strategic collaboration with research service and support providers to outsource support for "niche" solutions; be a trusted provider through associations with EOSC and its partners
    - increased revenue; reputation for serving the research market.

Role: Research Support Provider: for example, support staff at RI, support staff at e-Infrastructure, support staff at university (e.g. data steward, computing engineer)

- **Goals** may include providing efficient support to researchers within the institution, proactively seeking out new ways of working to support innovation, gaining more expertise
- **Potential gains from EOSC** include
    - minimising the effort to support new use cases; avoiding working on a blank canvas for new requests
    - finding new services and solutions to propose to their own supported researchers

- potential to provide new services based on new ways of working inside the EOSC
- enabling more efficient research; improving recognition of the value of support activities.

Role: Research Asset Producer: for example, researcher (as producer of new research outputs), research team

- **Goals** may include disseminating outputs more widely, achieving wider reuse of outputs, demonstrating practice of open science
- **Potential gains from EOSC** include
  - increased findability of their outputs through EOSC portal/catalogue
  - increased reuse of research outputs due to information provided in the EOSC catalogue (compliance with the Rules of Participation)
  - increased reputation and scientific impact by practicing open science

Role: Research Asset Consumer: for example, researcher (as consumer of existing research outputs), research performing organisation, institutions (IT departments etc), structured research communities, communities of practice, research engineer/software developer, citizen scientist

- **Goals** may include producing better research, easier access to research assets
- **Potential gains from EOSC** include
  - easier access to research assets
  - using EOSC to establish contacts with other researchers outside normal constituency
  - achieving higher quality research due to optimal choices of assets; producing open research outputs.

III.1.2 Key Relationships and Related Transactions

Once the roles had been identified, the value exchanged between them was mapped in a matrix. This information was used to prioritise three key relationships which were analysed in more detail. The aim of this exercise is to concentrate on interactions relevant for the scope under consideration (often those which are likely to exhibit the highest volume of transactions) which have the greatest potential for value to be derived from the platform. This involved interactions between service providers and research support providers, and between research asset consumers and research support providers. Figure 1 below shows the key relationships involved in the use cases. The relationship between research support providers and commercial service providers was not studied in detail because the focus of the work was on publicly funded services.

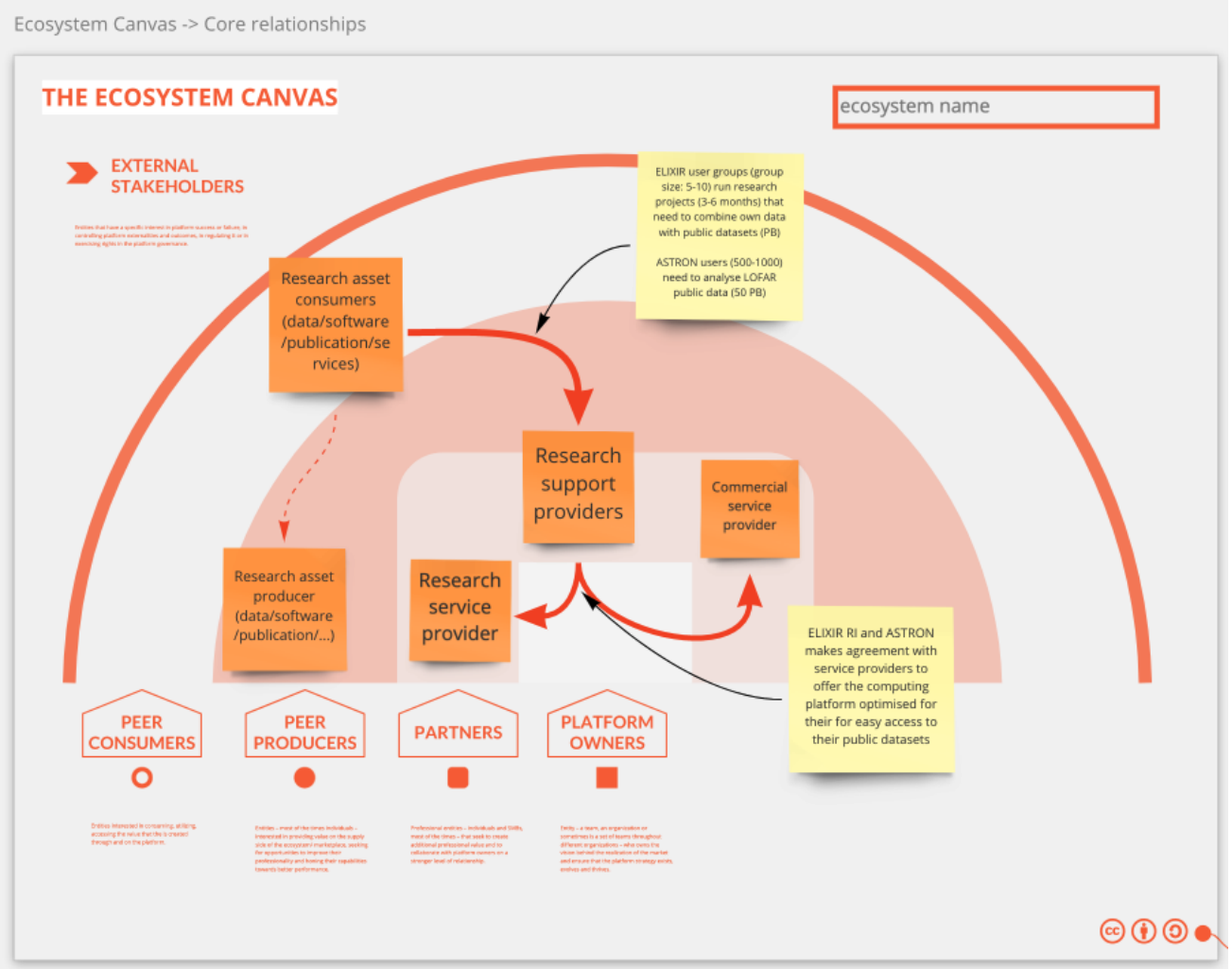

Figure 1: Key Relationships in the Use Cases

EOSC-HUB [illegible] FUNDING FROM THE EUROPEAN UNION'S HORIZON 2020 RESEARCH AND INNOVATION PROGRAMME UNDER GR[illegible]EMENT NO. 777536.

Once the key relationships were identified, transactions (i.e. elementary interactions among two roles) were analysed to understand: what is exchanged, the related value unit, and the channel/context. The goal was to identify improvements the platform strategy could bring to each transaction to make it *easier*, *cheaper* and *faster*.

For the core relationship <research asset consumer, research support provider>, the transactions identified were: request service, authorise service, use service, complete project.

For the core relationship <research support provider, research service provider>, the transactions identified were: request consultancy on science case, express interest in evaluating science case, confirm evaluation meeting, run consultation meeting, make proposal, accept offer and commission service, provision service.

The modelling identified aspects of these interactions where improvements could be made compared to the current situation:

- provide templates for use case description, to better extract use cases from the demand side to the supply side
- offer one-to-many interaction with providers (i.e. advertise demand), to allow a broader and more standardised response from suppliers to demanding parties
- standardise metadata, to enable a more standardised discovery of data, software and services and allow for easier comparison
- standardise access terms, to create a more consistent and predictable access and provision of services
- provide standard contracts, to bundle best practices and increase the professionalism in public sector
- provide information on funding solutions and conditions, to allow for better signposting and lower the threshold for low-capital research groups
- ease cross-border use of research assets such as storage and compute services, to support international organisations' research
- provide quality and usage statistics
- facilitate aggregated usage across wider set of researchers
- support networks of research support providers for knowledge exchange
- increase quality and availability of research support
- provide user guides and best practice templates
- simplify services – e.g. intuitive interfaces, to lower the threshold for upcoming research disciplines and lower the costs of training and support
- support interoperability of services for composability
- support rating of research assets and artefacts – user feedback.

Whilst several of these points are already recognised as gains the EOSC should deliver, the platform design approach helps to highlight the increased volume of transactions (i.e. use of the EOSC for matching user needs and supplier offers), and the greater and more efficient use of publicly funded research assets which can be expected to result from making these improvements. The work clearly

identified the importance of the role of Research Support Providers and the value they could add to the EOSC.

III.1.3 Evolving In The Ecosystem

After identifying the key roles in the ecosystem, the core relationships and their transactions, an essential part of a platform strategy is to identify the support and enabling services that allow entities to improve in their role or eventually evolve and embrace new opportunities. For each of the key roles, challenges and possible services were analysed in three stages: onboarding the platform, getting better in the current role or catching new opportunities.

For Research Support Providers

- Key challenges at onboarding the platform stage: finding suitable providers for the scientific cases they want to support, quantifying needs for these cases
    - Platform potential to address: consultancy, catalogue of services, publication of scientific challenges (demand-side advertising)
- Challenges getting better in the current role: standardise framework agreements
    - Platform potential to address: develop framework agreements
- Catching new opportunities: provide support outside traditional user community
    - Platform potential to support/encourage: promotion channel and matchmaking platform to find communities that need support

For Research Service Providers,

- Key challenges at onboarding the platform stage: Listing services, pricing services, promoting services to a wider audience
    - Platform potential to address: provide cost/pricing models, service descriptions, guided onboarding and reputation statistics (quality, usage, user satisfaction)
- Challenges getting better in the current role: provisioning for-pay services, leveraging local funding schemes for provisioning
    - Platform potential to address: funding scheme knowledge base, updates to regulations, guidelines and best practices, procurement support
- Catching new opportunities: maintaining in-house status whilst being able to respond to new opportunities
    - Platform potential to support/encourage: provide framework that allows preserving in-house status on a wider scale

For Research Asset Consumers,

- Key challenges at onboarding the platform stage: finding needed assets, estimating trustworthiness of assets
    - Platform potential to address: catalogue of services, feedback and community support
- Challenges getting better in the current role: integrating solutions, receiving invoices

    - Platform potential to address: consultancy solution architecting, brokering to (local) funding schemes
- Catching new opportunities: becoming a data provider, establishing collaborations to improve the existing software or develop new applications, stepping up as an expert supporter
    - Platform potential to support/encourage: supporting a human network, data lineage and ingest tools, dynamic results from EOSC observatory.

It can be seen that the EOSC has the potential to support each of these key roles to evolve in ways which allow them to offer more value, or to derive more value from the EOSC. Overall, the EOSC can support better science and more effective use of research assets.

III.1.4 Assembling The Platform Experience

The final step of the PDT methodology is to assemble the roles, the transactions and the learning opportunities into an experience that synthesises the core value proposition essential for the platform strategy. While several different experiences can be supported, in this case the focus remains related to the specific scope of this document. The experience also includes considerations about sustainability models.

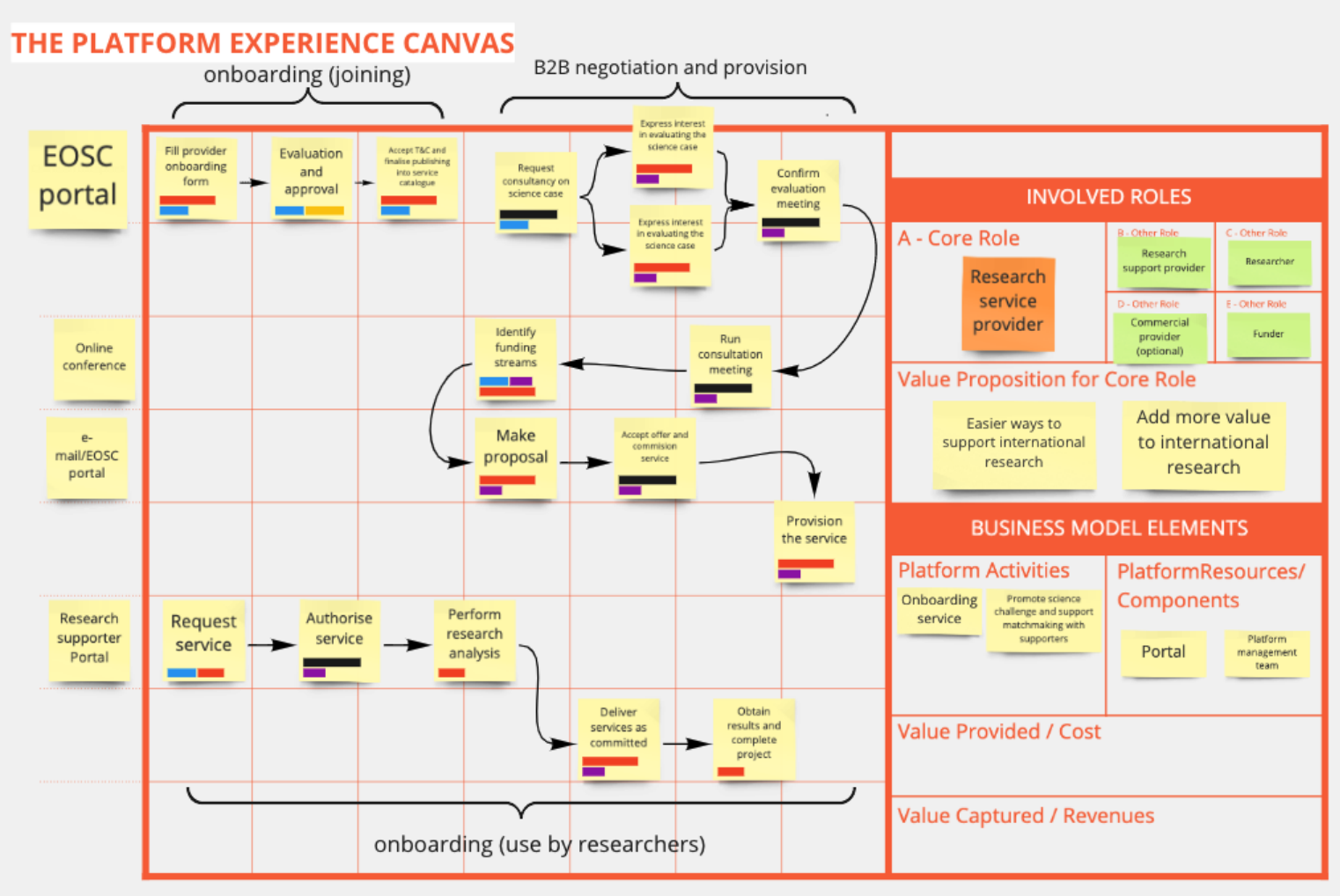

Figure 2: Matching of Demand and Supply

EOSC-HUB [illegible] FUNDING FROM THE EUROPEAN UNION'S HORIZON 2020 RESEARCH AND INNOVATION PROGRAMME UNDER GR [illegible] EMENT NO. 777536.

The current EOSC portal supports a platform experience where providers can onboard their services, while users can browse the catalogue and request them, either directly via the portal or indirectly via the providers' websites. This type of experience works well when the user has a clear understanding of what they want to achieve and is able to identify the needed service.

In the scientific computing sector, it is common that the user needs guidance and support to make optimal use of the available resources and technology. This need was highlighted by the platform design analysis, with the role of Research Support Provider as one of the primary beneficiaries of the EOSC portal. This fact is even stronger in the business-to-business scenario, where research organisations (Research Infrastructures and e-Infrastructures) need to discuss complex and large-scale needs as demonstrated by the two use cases presented in this document.

The canvas shown in Figure 2 depicts three main types of interaction with the EOSC: 1) providers onboarding their services in the EOSC portal (labelled onboarding (joining) in figure 2); 2) research organisation announcing a research challenge, receiving invitation to negotiations, identifying a solution, commissioning it and obtaining the service (labelled B2B negotiation and provision); 3) the users from the research organisation consuming the provisioned service for the community (labelled onboarding (use by researchers)). While this type of interaction is supported by the EOSC-hub project for specific cases (e.g. see Thematic Services and Competence Centre), it is not yet a capability offered by the EOSC Portal. The modelled "platform experience" highlighted the importance of identifying a funding source to pay for the service usage and also of understanding the incentives for Research Support Providers to operate in the EOSC ecosystem.

III.2 Further Information

The platform design toolkit is provided by Boundaryless. Their website, https://platformdesigntoolkit.com/, contains a large amount of information describing and illustrating platform design and the methodology for applying the platform design approach in business modelling. A few resources which may be of interest are highlighted below.

A good example of platform design:

https://stories.platformdesigntoolkit.com/pdt-bootcamp-example-explained-34e08f9dd4b3

A recap on platform design:

https://stories.platformdesigntoolkit.com/all-you-need-to-know-about-platform-design-in-a-handy-recap-6a1d1d9b5de0

Platform design stories (user orientation):

https://stories.platformdesigntoolkit.com/navigating-platform-design-toolkit-f0f39cdf557

Example of use of platform design in public policy context (UNDP):

https://stories.platformdesigntoolkit.com/mobilizing-ecosystems-towards-sustainable-development-goals-with-platform-strategies-ccbd9e494ba9

Public policy making as a platform:

https://stories.platformdesigntoolkit.com/public-policy-making-as-a-platform-b49f2549a2b3

Glossary

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

Terminology/Acronym	*Definition*
ASTRON	The Netherlands Institute for Radio Astronomy[72], is part of the institutes organisation of NWO
CESNET	An association of universities[73] of the Czech Republic and the Czech Academy of Sciences
Container	A software package that contains everything the software needs to run
CPU	Central processing unit, indicates a processor
CSC	A non-profit state enterprise[74] in Finland which develops, integrates and provides high-quality information technology services and ensures that Finland remains at the forefront of development
EGI Federation	A federation of computing and storage resource providers united by a mission to support research and innovation with advanced computing services. The federation is governed by the participants represented in the EGI Council and coordinated by the EGI Foundation.

[72] https://www.astron.nl
[73] https://www.cesnet.cz/cesnet/?lang=en
[74] https://www.csc.fi/en/home

ELIXIR	An intergovernmental organisation[75] that brings together life science resources from across Europe
EMBL-EBI	The European Bioinformatics Institute[76] is part of EMBL, Europe's flagship laboratory for the life sciences
EOSC	The European Open Science Cloud promoted by the European Commission to provide all researchers, innovators, companies and citizens with seamless access to an open-by-default, efficient and cross-disciplinary environment for storing, accessing, reusing data, tools, publications and any EOSC Resource for research, innovation and educational purposes[77].
EOSC Executive Board	Body of representatives from the research and e-infrastructures communities, appointed by the European Commission[78]
EOSC Governance	Overall Governance Structure for EOSC, comprising EOSC Governance Board, EOSC Executive Board and Stakeholder Forum[79]
EOSC Governance Board	Also "EOSC board": institutional group gathering the member states and the Commission to ensure effective supervision of the implementation of the EOSC[80]
EOSC-hub	Project creating the integration and management system of the future European Open Science Cloud[81]
EUDAT CDI	European e-infrastructure of integrated data services and resources to support research[82]
FAIR	Guiding principles to make data Findable, Accessible, Interoperable, and Reusable

[75] https://elixir-europe.org
[76] https://www.ebi.ac.uk
[77] https://eoscpilot.eu/eosc-glossary#overlay-context=eosc-glossary
[78] https://www.eoscsecretariat.eu/eosc-governance/eosc-executive-board
[79] https://www.eoscsecretariat.eu/eosc-governance
[80] https://www.eoscsecretariat.eu/eosc-governance/eosc-governance-board
[81] https://www.eosc-hub.eu/
[82] https://eudat.eu/

Federating Core	The activities, policies and resources required to facilitate, monitor and regulate as appropriate day-to-day transactions across the EOSC
GÉANT	Pan-European research and education network that interconnects Europe's National Research and Education Networks (NRENs)[83]
GPU	Graphics Processing Unit, suitable for parallel processing
Horizon 2020	The European Union Framework Programme for Research and Innovation
HPC	High Performance Computing
HTC	High Throughput Computing
INDIGO-DataCloud	Project developing a data/computing platform targeted at scientific communities, deployable on multiple hardware, and provisioned over hybrid (private or public) e-infrastructures
INFN	The National Institute for Nuclear Physics[84] is the Italian research agency dedicated to the study of the fundamental constituents of matter and the laws that govern them, under the supervision of the Italian Ministry of Universities and Research (MUR)
LOFAR	Low Frequency Array telescope experiment[85]
MPI	Message Passing Interface, a standardized and portable message-passing standard designed to function on a wide variety of parallel computing architectures
NWO	The Dutch Research Council[86] which funds top researchers, steers the course of Dutch science by means of research programmes and by managing the national knowledge infrastructure
OpenAIRE-Advance	Project supporting Open Access/Open Data mandates in Europe

[83] https://www.geant.org/
[84] www.infn.it
[85] http://www.lofar.org
[86] https://www.nwo.nl/en

Shared Resources	Resources including scientific outputs (local copies of data; applications, software, pipelines etc.) and the storage and compute hosting platforms needed to deposit, share and process them. The shared resources realise the EOSC "**resource tier**".
SURF	The collaborative organisation[87] for ICT in Dutch education and research
VM	Virtual Machine, a virtualised server environment on which a guest operating system and associated applications can run

[87] https://www.surf.nl/en