D2.4:The EOSC Delivery Chain

Author(s)	Lars Fischer, Per-Olov Hammargren, Leah Riungu-Kalliosaari, Maijastiina Arvola (CSC)
Status	Final version
Version	1.0
Date	14.08.2020

Document identifier:	
Deliverable lead	NORDUnet
Related work package	WP2
Author(s)	Lars Fischer, Per-Olov Hammargren, Leah Riungu-Kalliosaari, Maijastiina Arvola
Contributor(s)	
Due date	31.08.2020
Actual submission date	15.09.2020
Reviewed by	Francesca Iozzi, Valter Nord
Approved by	
Dissemination level	Public
Website	https://www.eosc-nordic.eu
Call	H2020-INFRAEOSC-2018-3
Project Number	857652
Start date of Project	01/09/2019
Duration	36 months
License	Creative Commons CC-BY 4.0
Keywords	Delivery chain, Cross-border collaboration models, ICT services

EOSC-Nordic project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 857652

EOSC-Nordic project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 857652

Executive summary

This deliverable is the first of a series of three studies of cross-border service delivery and collaboration in EOSC planned by EOSC-Nordic. We look at EOSC from the perspective of the delivery chain and identify the different actors that belong to the EOSC delivery chain and map their roles and responsibilities, based on a number of service case studies. Future work will study existing cross-border collaboration models in the Nordic and Baltic countries (for publication in February 2021), and propose recommendations for future cross-border collaboration in the context of EOSC (for publication in August 2021).

A successful delivery chain involves many actors. Each actor plays a critical and unique role that is crucial to the delivery of the services. Becoming a part of the EOSC service ecosystem means dealing with additional layer(s) of complexities and interconnections in the delivery chain, which is an aspect that should be planned and executed meticulously.

The work reported here is part of the *Policies, legal issues and sustainability* effort of the EOSC-Nordic project to foster coordination of and connection between national initiatives at policy level, in the context of EOSC. The deliverable will help policy makers stay informed about cross-border service delivery and service-provision in the context of EOSC. The deliverable will also be input to the EOSC governance and architecture evolution.

In this report, we consider service case studies and illustrate that the service delivery chain for ICT services for research requires complex and careful technical and organisational orchestration, dependent on the interplay of many stakeholders, on existing national and international agreements outlining rights and responsibilities, and on tradition and human networks. We show that these relationships depend on the nature of the services, the community of users, and the basis for the funding of the services. We have seen also that these relations are especially complex in a cross-border service delivery environment.

We analyse different aspects of the service delivery chain, considering user communities, funding, operations and support, resource types and governance. We show that all these aspects influence how the service may be offered and its relation to EOSC. We show that this takes many forms and is different for all of the six service cases considered. The cases have been chosen on the basis that they all reveal important aspects of the delivery chain, and for this reason, some of the cases have an above-average delivery chain complexity.

Our analysis of the six services reveals that some of these services are generic and can be offered universally through EOSC, provided that funding to replenish consumable resources can be ensured. Other services that mainly provide access to data can be provided through EOSC, without additional funding, when the stakeholders have incentives to do so. Some services are national in scope and governance and may therefore be regarded as strategically important nationally. For such services, there might be policy considerations to take into account - in addition to the technical and workflow challenges. Finally, some services are specifically tailored towards a research community and as such integral to the research practices of that community. Incentives for the community to offer these services through EOSC may not be clear and should be assessed case by case.

EOSC-Nordic project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 857652

EOSC NORDIC

Table of Contents

EOSC-Nordic project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 857652

1. Introduction

The purpose of this document is to investigate the delivery chain for services in light of the European Open Science Cloud (EOSC), to identify the actors involved in the service delivery and map their roles and responsibilities, and to analyse the implication for inclusion in the EOSC framework. The investigation is based on existing services in the Nordic and Baltic region with the aim of highlighting the service delivery complexities.

The future EOSC aims to enable service provisioning of broad and secure access to data resources and research data, processing services, as well as foster technical development, and managing communications with researchers, communities, and service operators, via a federated virtual environment[1]. In order to participate, researchers, communities, and organisations delivering services are expected to make services available via federated EOSC solutions, such as the EOSC-Exchange[2], in accordance with the EOSC Rules of Participation (RoP), any terms stipulated within forthcoming EOSC statutes, as well as the EOSC partnership. The terms and conditions will include aspects related to data sharing, services, rules for operators, and general terms and conditions.[3]

Delivering services for research is a complex and multifaceted process, involving many stakeholders. Service delivery requires funding, governance, operations, maintenance, and engagement activities with the scientific communities, as well as support. This deliverable aims to illuminate these aspects of the service delivery chain, and their implications for EOSC.

As such, this deliverable focuses on EOSC from the perspective of the delivery chain. Delivery entails more than a relationship between a service provider and an end-user. Many actors are present in the delivery chain, working as regulators, intermediaries or providing support and consultancy work for the actual delivery of resources, such as e-infrastructures, research infrastructures, universities, research and education networks, etc. This is especially the case when complex services are being delivered. In a cross-border environment, this chain becomes even more complex[4], involving actors from different countries operating from different administrative and/or scientific domains and organisations.

In order to put the aspects related to the delivery chain in context, we selected six existing services with the aim of highlighting the service delivery complexities. For this reason, cases that illustrate the complexity of the service delivery chain were chosen. Not all services are as complex as the ones demonstrated in this deliverable. Four services are provided by national e-infrastructure providers in Estonia, Finland, and Sweden. The remaining two cases provide an overview of service provision within thematic research infrastructures, specifically ELIXIR and ICOS ERIC.

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

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

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

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

EOSC-Nordic project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 857652

2. EOSC and Cross-Border Delivery Chains

When a service is invoked, e.g., by being requested in an EOSC portal, a chain of events must take place to deliver the service to the user and allow the service to be used. The chain may start before the service is invoked (i.e., contracting, service level and privacy agreements), and may have elements that continue after service use (i.e., accounting, payment). The chain has legal, technical, and governance elements. We call this process the *service delivery chain*.

In logistics for physical products, we talk about the supply chain. The supply chain involves component suppliers, transportation of partial and finished products, warehouses, the management of the flow and timing of physical objects. It also involves collaboration of multiple organisations, sharing of inventory information, contracting, rights, obligations, and payments. This frequently involves intimate collaboration across organisational boundaries when new products are developed or the manufacture and supply of existing products changes. This process is highly complicated and critical to the delivery of physical products.

Likewise, the delivery chain for ICT services for research involves many stakeholders. The EOSC Rules of Participation mention service providers, operators, and users[5]. In EOSC, these may belong to different organisations and be in different countries. The service delivery chain for an ICT service is the relation between these stakeholders. The service delivery chain ensures that services are operated and maintained, and supports the evolution of the service in accordance with the needs of research communities. Crucially, the delivery chain is a process that serves to align the stakeholder organisations to ensure each actor understands and fulfils their roles and responsibilities. This is especially critical for cross-border collaborations, where services must be delivered, operated, and maintained by partners in several countries and serve a user community distributed across several institutions and countries.

2.1.1 Actors, Roles, and Responsibilities

To make the service delivery chain work, a number of actors must participate, taking on different roles and responsibilities. EOSC will serve science across multiple institutions and multiple countries, and so the actors, roles and responsibilities for EOSC services must be effective in a cross-border setting. This has impact of every step from funding and governance to operations. Furthermore, roles and responsibilities of the actors must cover the entire lifecycle of research data, from creation to sharing, publishing and archiving.

We selected a limited number of services to illustrate the different actors, roles and responsibilities across various organisations. An overview of this can be found in the appendices.

In Chapter 3 we briefly present the service cases. In particular, we focus on the purpose, usage and management of each service to illustrate the width of the service delivery chain, from forming cross-border collaboration and commitment to funding, to operations management of the service.

In chapter 4 we discuss the governance of services. We look at governance mainly from the perspective of national institutions and priorities, and to a lesser extent from the perspective of international research collaborations and research infrastructure.

In chapter 5 we then offer an analysis of a number of relevant aspects of the service delivery chain, based on the examples presented in chapter 3 and 4.

[5] https://repository.eoscsecretariat.eu/index.php/s/QWd7tZ7xSWJsesn#pdfviewer

EOSC-Nordic project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 857652

3. Service Case Studies

In order to put the aspects related to the service delivery chain in context, we selected six existing services with the aim to highlight the complexities of service delivery. The service cases will give the reader a hands-on understanding of the service delivery chain for ICT services for research.

Four of the cases are services provided by national e-infrastructure providers, while two cases are provisioned within thematic research infrastructures. They are not meant to be comprehensive across the geographic area, service types, or scientific disciplines. Rather, the services are picked to illuminate the range of complexity of service delivery.

The service cases are all existing or potential EOSC services. The national services are examples of services that can benefit users in other member states by being included in EOSC (and if national policy allows this). The ICOS services are already part of the EOSC-hub project, and ELIXIR is a pan-European, thematic Data Infrastructure, which in many ways resembles EOSC within a specific field of research.

The services selected have a user community in the Nordic and Baltic region and are provided either by service providers of the region or by Research Infrastructures with funding and operations participation in the region. Furthermore, members of the EOSC-Nordic project are directly involved with the services and Research Infrastructures. Taking advantage of this, the case descriptions have been developed by subject matter experts from the EOSC-Nordic partners, based on hands-on experience with governance and operations of the services as well as engagement with the user communities.

The service cases are as follows:

ELIXIR Bioinformatic RI: ELIXIR is a biological information infrastructure in Europe as well as one of the most important infrastructures globally, connecting scientists across the globe with a vast amount of bioinformatics data. ELIXIR is a distributed research infrastructure, operated by a number of national nodes and service providers, and dependent on a range of bi-lateral and international agreements among the participating countries and institutions. ELIXIR is an example of community-focused services, established and governed as part of a discipline-specific enterprise.

UT Rocket: UT Rocket is a High-Performance Computing (HPC) service provided and operated by the Estonian Scientific Computing Infrastructure (ETAIS) at University of Tartu in Estonia. It is an example of an institutionally operated service that is available nationally for a broad range of researchers, across disciplines based on national funding.

IDA: IDA is a Finnish national service for continuous and safe storage of research data provided and operated by CSC – IT Centre for Science. It is part of a FAIR data family of services providing integrated solutions for storing, sharing and publishing research data. It is an example of a nationally funded service established centrally to further the national objective of FAIR data, across all scientific disciplines.

SNIC Swestore dCache: Swestore dCache is a distributed storage service supporting active (non-sensitive) data with low to moderate storage needs, which is provided through the Swedish National Infrastructure for computing (SNIC). It is an example of a federated service operated locally but funded and offered nationally, with access granted by scientific and technical review.

SNIC Alvis: Alvis is a compute resource for Artificial Intelligence/ Machine Learning (AI/ML) research, providing unique technology for AI/ML approaches to scientific data analysis. It is an example of a specialised ICT service provided and operated by SNIC for a specific methodology, with the objective of

EOSC-Nordic project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 857652

furthering this methodology nationally. It is furthermore an example of a service funded by a public foundation, with objectives set out by a multitude of stakeholders, including the private funder.

ICOS ERIC: The Integrated Carbon Observation System, ICOS operates in three domains, namely atmosphere, ecosystem and ocean[6]. It consolidates open data from more than 130 measurement stations across 12 European countries. It is an example of services that exist primarily to make datasets available and accessible. It is also an example of a collaborative service provided by a European Research Infrastructure Consortium (ERIC).

More details on the services are available in the appendix section. Appendix A contains tables illustrating the actors involved in the service and their corresponding roles and responsibilities. Appendix B provides a detailed description of the service and its place within the e-infrastructures ecosystem.

4. Service Governance

Services, such as those to be made available via EOSC-Exchange, are provided by a range of organisations governed by mandates in steering documents. Service governance is defined as mandates in steering documents. The steering documents described are directives, instructions and such, from governments, funders, and other types of public funding bodies, which details mandates for organizations.

The aim of this chapter is to exemplify governance of services such as networks provided by National Research and Education Networks (NREN), and compute, storage, provided by national e-infrastructure providers funded by national funds, which have a national mandate in the Nordic and Baltic region. The chapter also describes governance of services delivered by ESFRI research infrastructures[7] and ERICs.

As described in the EOSC Partnership proposal, service delivery of the future EOSC is in part to rely on in-kind-contributions from members of the EOSC-Association[8]. Therefore, it is important to understand different national models of service governance.

4.1 Sweden

In terms of service governance, the mandate of SNIC is set in the Swedish Research Council specific grant agreement (GA) and the SNIC consortium agreement (CA). The GA sets the mandate for SNIC in terms of the delimitation of the user base of SNIC resources. SNIC makes available resources for researchers affiliated with Swedish Higher Educational Institutions (HEI), as well as other Swedish authorities with research assignments included in the HEI concept.

SNIC activities, including service delivery, are governed via third-party agreements between the Uppsala university, as the host organisation for SNIC, and on other hand the consortium parties. SNIC is not a legal entity of its own. The owner of a SNIC resource, a party of the consortium, is legally responsible for that particular resource. Ownership of data typically belongs to the HEI with which the respective Principal Investigator (PI) is affiliated.

[6] ICOS in a nutshell: https://www.icos-cp.eu/about/icos-in-nutshell

[7] European Strategy Forum on Research Infrastructures Roadmap 2018. https://ec.europa.eu/info/sites/info/files/research_and_innovation/esfri-roadmap-2018.pdf

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

EOSC-Nordic project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 857652

The mandate of SNIC is as such set out in two agreements (GA, CA), and actual service delivery is stipulated via third party agreements. Potential service delivery through EOSC channels, such as the EOSC Exchange, would as such entail three interdependent layers of service governance having to be adjusted to facilitate the actual delivery of the service.

4.2 Denmark

The establishment of the Danish e-Infrastructure Cooperation (DeIC) became official in 2012, via a legal document signed by the Minister of Science and approved in parliament. This document provided the legal vehicle for setting up a formal organization and authorized the Ministry to allocate resources and mandate to the organisation.[9]

Services within DeIC are governed by the General agreement between the Ministry of Science and the Universities, that defines governance, user base and general ambition for the national IT provisioning as well as funding and access principles.[10] In order to bridge the transition, all the universities have agreed to provide access to resources locally on a pay-per-use principle.

The formal governance of DeIC is established through statutes lastly modified 2019.[11] Since the main part of DeIC's staff is formally employed by the Danish Technical University (DTU), there is a formal agreement between DTU and the Danish Conference Vice Chancellors establishing the responsibilities of DTU in providing human resources to DeIC and the responsibility of the Universities to fund that activity.

4.3 Latvia

The Latvian Ministry of Education and Science (IZM) is responsible for maintaining functionality and governance of the **Latvian NREN**.[12] IZM is the Latvian member in GÉANT, but the connection of Latvian academic network LAT to GEANT has been delegated to be provided by SigmaNet - the Academic Network Laboratory of the University of Latvia. IZM provides funding for the academic network LAT and the Data Centre facilities. Access to LAT is open and free of charge for academic and research organizations for educational and research (non-commercial) purposes. Funding of research institutions, including related e-infrastructure and databases, is regulated by the Order No.1316[13].

There is no organized national level HPC and data services, but academic and research institutions can place computing and/or data storage equipment in the Data Centre facility. This process is coordinated by IZM. Academic and research organizations are organizing these services by themselves. Local network infrastructures of academic institutions, as well as computing/data equipment and services, are in responsibility of their owners and are financed from available resources (government donations, user contributions, EU project funding etc.).

Each computing service provider has its own access policy and application site for new users. Infrastructure is open mainly for academic users, and industry users for research purposes. There are regulations not allowing usage of LAT for other than academic organizations (ex. schools) and for providing commercial

[9] https://www.deic.dk/sites/default/files/u50/PDF/Aktstykke_Finansudvalget_190420012.pdf

[10] https://www.deic.dk/sites/default/files/u50/PDF/Aftale_af_30_september_2011_om_reorgansiring_og_styrkelse_af_dansk_e-Science_DOK2044078%20%282%29.pdf

[11] https://www.deic.dk/sites/default/files/uploads/PDF/DeiC%20Vedt%C3%A6gter%20af%2025-11-2019.pdf

[12] https://likumi.lv/ta/id/273100 (in Latvian)

[13] https://likumi.lv/ta/en/en/id/262508

EOSC-Nordic project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 857652

services. Research institutions involved in ESFRI projects use the LAT network, their own compute/data equipment, as well as other services available on mutual agreements.

4.4 Finland

CSC – IT Center for Science is a non-profit state enterprise with special tasks. As part of the national research system, it develops, integrates and provides information technology services. CSC is owned by the Finnish state (70%) and higher education institutions (30%). Due to the ownership model, CSC is also governed by the State Shareholdings and Ownership Steering Act[14]. CSC is organized as a limited liability company with the according legal requirements and obligations (Finnish Limited Liability Companies Act 624/2006)[15] reflected in the governance model.

CSC's primary customers are the Ministry of Education and Culture and organizations in the field, higher education institutions, research institutes and public administration. CSC's owners are entitled to buy services from CSC without a competitive call included in the procurement, and services are offered by CSC on a non-profit basis. Service governance is thus implemented also through these contracts and agreements between CSC and its customers/owners.

In addition, Finland's Strategy and Roadmap for Research Infrastructures 2014–2020[16], includes Finland's research infrastructure strategy and an updated roadmap for 2014–2020. The research infrastructure ecosystem is governed by this strategy. Such an ecosystem includes major national research infrastructures, Finnish actors' partnerships within European infrastructure projects (ESFRI), memberships of other international infrastructures and research organisations' strategically significant infrastructures. In 2020, the Academy of Finland updated a strategy for national research infrastructures in Finland for the years 2020-2030[17].

4.5 Estonia

The Ministry of Education and Research in Estonia implements the national research policy, organises the financing and evaluation of the activities of R&D institutions and coordinates international research cooperation at the national level. The Ministry is also responsible for the planning, coordination, execution and monitoring of research policy related to the activities of universities and research institutes[18].

Estonian Research Council is in charge of coordination of the implementation of the set goals, including arrangement of the Open Science Expert Group sessions. The latter is tasked to elaborate proposals for developing the principles of national open science policy and to advise Estonian Research Council as well as the Ministry of Education and Research and other related parties on open science issues (both strategic and practical), like: collecting scientific information, open access publishing, infrastructure pertaining open science and its data etc.

Estonian Scientific Computing Infrastructure (ETAIS) belongs to the Estonian roadmap of research infrastructures providing computing and storage resources for Estonian scientific community. The ETAIS

[14] https://www.finlex.fi/fi/laki/kaannokset/2007/en20071368.pdf

[15] https://www.finlex.fi/fi/laki/kaannokset/2006/en20060624.pdf

[16] https://www.aka.fi/globalassets/awanhat/documents/firi/tutkimusinfrastruktuurien_strategia_ja_tiekartta_2014_en.pdf

[17] https://www.aka.fi/globalassets/42julkaisut/aka_tik_strategia_2019_en_digi_a.pdf

[18] https://www.hm.ee/sites/default/files/estonian_rdi_strategy_2014-2020.pdf

EOSC-Nordic project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 857652

project is being carried out by a consortium of four institutions: University of Tartu, Tallinn University of Technology, National Institute of Chemical Physics and Biophysics and Information Technology Foundation for Education.

4.6 Norway

UNINETT Sigma2 AS (hereafter called Sigma2) was established in 2014 under the national agreement between the Norwegian Council of Research (NCR) and the Metacenter consortium consisting of the Norwegian four oldest universities. The scope of Sigma2 is to optimize the impact of research by providing a sustainable, predictable and cost-efficient e-Infrastructure. The service landscape procured and operated by Sigma2 consists of High-performance computing facilities, research data infrastructure, on-demand computing resources, data services and data archiving facilities. The operations are run in collaboration with the local IT-departments at the Metacenter sites. From this perspective the metacenter sites have a dual role of funding stakeholders and co-operators in the governance of the infrastructure.

The basis funding for operations and part of the procurements come from the NCR and the consortium partners, while the large part of the investments is covered by resources granted by the NRC on a competition level every second year (through the call for INFRASTRUKTUR program). Services are available to individuals and groups involved in research or education at Norwegian universities, colleges and other research organizations. The applied cost model differentiates the price according to the service delivery mechanism. The majority of the resources are delivered through evaluations of the scientific merit and minimum costs are associated to the granted resources, while the dedicated resources have higher price but are not subjected to scientific revision.

Sigma2 is mandated toward the delivery of services for massive data and computing resources. Other services required to cover the needs of the long tail of the research production shall be covered by the local institutions and delivered through other locally defined delivery chain.

4.7 Thematic Research Infrastructures

Thematic Research Infrastructures are funded by both countries and agencies, and often have national and European institutions as members, with the aim to support a specific research discipline or disciplines. The governance of thematic research infrastructures service has the dual goal of ensuring agreements and support among entities funding the infrastructure and ensuring representation for the research community using the infrastructure.

RIs are governed by grant agreements and consortium agreements, setting the scope for both use of finances and activities of the RI. These agreements may stipulate governance boards, scientific advisory boards, resource allocation committees, and so forth. The description of the ICOS ERIC and the ELIXIR RI in Appendix B give examples of the complexity of governance of a multi-country infrastructures, as well as the different ways in which stakeholder representation can be ensured.

4.8 Summary

As exemplified above, services are governed in a variety of ways throughout the Nordics and Baltics. In some cases, governance is an element of a series of agreements between a multitude of stakeholder institutions, in some cases between one institution and one or more government agencies. In all cases, governance serves to define responsibilities and to prioritise resources. This can take the form of deciding

EOSC-Nordic project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 857652

who can use limited resources, prioritising the evolution of the service, or prioritising the use of stakeholder resources to support the services.

5. Analysis

This chapter presents an analysis of the service cases. Based on aspects of the service delivery chain, the analysis highlights a range of challenges and issues that should feed into the discussion on EOSC service delivery chain.

A central observation for the cases investigated here is that many of the services have complex stakeholder relationships that include provider and user, but often also funders, policy makers, community members, etc. Often, these stakeholder relations are interrelated. Below we illustrate aspects of these complex relationships, and how the actors take on different roles and responsibilities to orchestrate the service. We observe that this orchestration is broader than the service delivery, starting before the service is requested and extending past an instance of service use, and that each of the elements in this service delivery chain are critical.

5.1 Aspects of the Service Delivery Chain

To illustrate the complexity of service delivery, and how different aspects impact the delivery chain, we have reviewed a number of important aspects of the service delivery chain. The objective here is to highlight the relationship between services, service providers, and EOSC. This could in turn offer input towards guidance for service on-boarding.

Below, we will review the service delivery chain in terms of user community, funding, operations and support, resource types, and governance.

EOSC-Nordic project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 857652

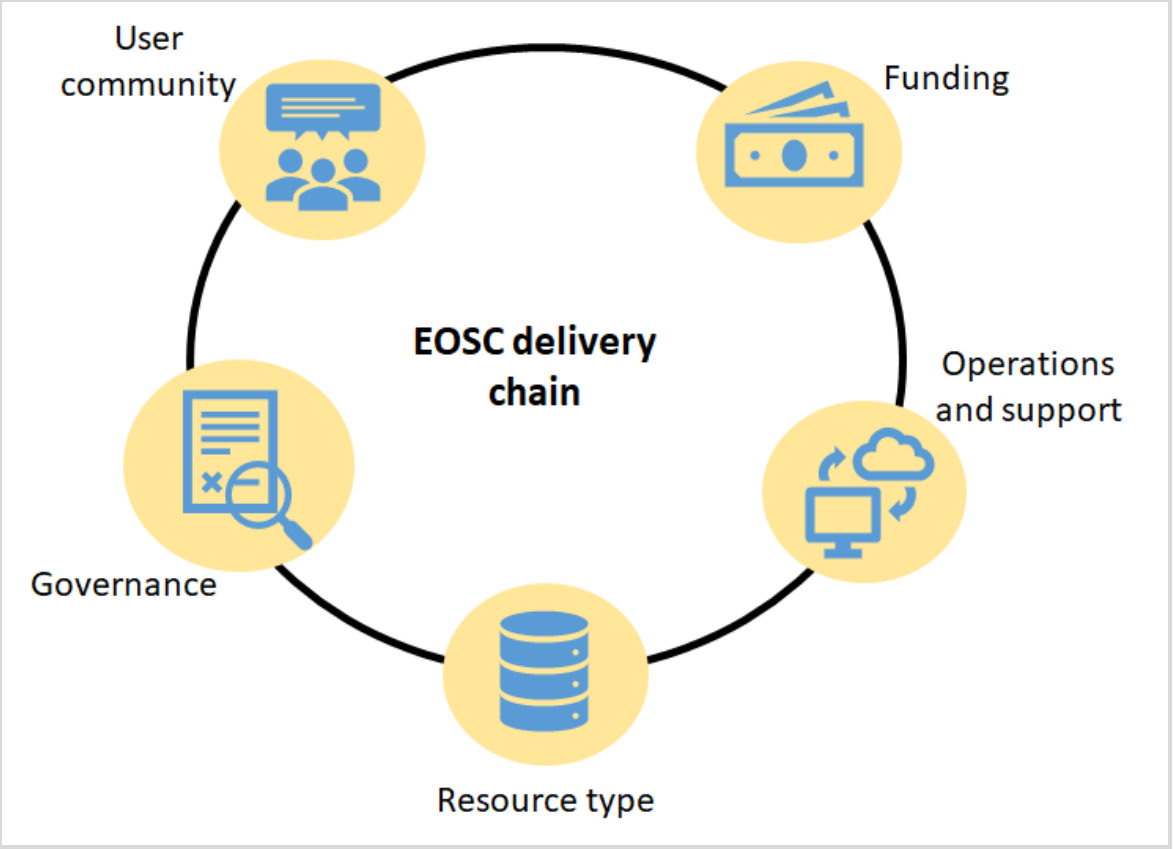

Figure 1: EOSC service delivery chain

5.2 User Community

Services are in general delivered with a specific community of users in mind - often within a certain field of research and/or geography. In some cases, the service may be engineered to solve a narrow problem that exists only for a specific community of users. In other cases, the service may be generic, but integrated into other services of a community or hosted close to the users of that community.

In the case of *UT Rocket*, the user community is any user with computing needs not requiring HPC scale facilities. The service can in general be delivered to users anywhere, provided a business model for the transaction can be established.

In the case of *Swestore*, the user community is restricted to users from national (Swedish) research institutions. This assumption of national service delivery is built into how the service is hosted and managed (distributed across national resource centres), how users are supported (by the SNIC organisation), how it is governed, and how it is funded. In this way there is an overlap between the service delivery chain and stakeholder management, with user communities, funders and service provides being represented. In the case of Alvis, the users are those who advance the application of a specific technology in science, advancing the role of this technology in a specific country, and thereby the standing of this country in this technology area. The service is funded by a private entity, specifically to serve Swedish users.

In the case of services in ELIXIR and ICOS ERIC, a number of services, small and large, are provided for a global community of users within a specific science discipline. The services are integrated to support each other and are delivered from a number of institutions that are all engaged with the user community of

EOSC-Nordic project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 857652

choice. As such ELIXIR can be described as a service web rather than a service chain, not being dependent on the single links in the chain to deliver service.

Some of the services in ELIXIR and ICOS ERIC are engineered to meet specific requirements of the user communities, while other services are generic in nature. Community services are integrated with other services in ways that ensure that they serve the community as a whole. The service portfolio includes a number of discovery services that allow users in the community to find the resources they need. The needs of the specific user community are built into nearly every other aspect of the service delivery - how and where the services are operated, the international agreements on funding, data sharing and service use, how the stakeholders fund the services, how the stakeholder govern the evolution of both individual services and the comprehensive service portfolio, etc.

For the suite of ELIXIR and ICOS ERIC services, access and security through user support to service hosting and funding are all part of a complex set of agreements between multiple stakeholders, entered into with the particular user community in mind. The stakeholder arrangements also influence how services are discovered and supported. Some services (e.g. data storage) may be generic and could be used by users outside the intended community, but such users would lack the relation to the other stakeholders.

In the case of IDA, the primary user community is Finnish researchers. The service is offered to research activities based at research institutions in Finland, for long-term archival of Finnish research output, and in order to increase the impact of Finnish research. The service is general and, in this sense could be re-targeted to similar objectives for other countries.

In all cases, the target user group influences how the service is delivered and distributed. It influences the evolution of the service, who will engage in supporting the service both technically and funding wise. It influences how users discover and use the service. In short, all stages of the service delivery chain are influenced by the user targets. For services to be delivered to another community of users, a range of assumptions, often built into the very fabric of the service delivery chain, need to be revisited.

5.3 Service Funding

Funding of services for research and education in Europe have a significant role in service evolution and influence most aspects of the service delivery chain. How funding streams are orchestrated significantly impacts service delivery, and thus the potential expansion into EOSC.

Funding is often based on multi-year agreements between various stakeholders. These agreements may come in the form of research grants, national or international grants to further a specific scientific discipline, or as part of a long-term or semi-permanent commitment to base infrastructure, either nationally or internationally.

The UT Rocket service is funded by the UT HPC facility, as part of a national agreement (ETAIS) where each partner contributes resources to a common infrastructure. The procurement of UT Rocket facilities has been supported by the EC. As UT Rocket funding is part of the ETAIS partnership agreement, partners have an expectation to have access to resources – and funders expect to get a proportional share of e.g. HPC resources linked to their individual funding level.

IDA is funded by the Finnish Ministry of Education and Culture, with a grant to the national Finnish e-Infrastructure operator CSC. The funding is provided to enable open access to Finnish research data.

Swestore, as other SNIC services, is funded by the Swedish HPC consortium SNIC, specifically for use by Swedish researchers. The funding is part of a long-term agreement between the Swedish Research Council, SNIC, and SNIC member institutions, and represents a Swedish national commitment to a storage facility for research data. The partnership is specifically national, and the funding is part of the national strategy for research infrastructure.

EOSC-Nordic project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 857652

SNIC Alvis is funded by the Wallenberg Foundation[19], as one of several grants by the foundation in support of Swedish research. The funding is specifically targeted at advancing Swedish AI research and methodology and is built into the governance of the service.

ELIXIR is a distributed infrastructure, relying on a federating national funding. The hub is funded through membership fees and Horizon 2020 project grants. This funding is channelled back to the ELIXIR nodes, which may also be financially supported nationally or by private foundations. The funding relies on agreements between a large number of stakeholders, who each fund part of the e-Infrastructures under an assumption of reciprocity.

Similarly, ICOS ERIC is funded through membership contributions from countries. ICOS ERIC also seeks finance from the EU and the national facilities from their countries' national resources. The funding is heavily dependent on the country's support to the RI.

The service funding is an integral part of the service delivery chain. Not only is the funding required to establish the service, it is essential to ensure that the service is sustainable over time. Unless a service can be sustained, science cannot rely on that service. This is especially true for handling, storing, and making available and findable research data, the core of the EOSC mission. The continued availability of funding for a sustainable service relies not just on one-time grants given before the service is created. It relies also on continuous feedback to/from stakeholders to ensure commitment to the service. This feedback is an integral part of the service delivery chain. This feedback happens in many ways, and relies often on national communities, discipline-specific collaborations, or explicit restrictions imposed by funders.

5.4 Free at the point of use

The EOSC Rules of Participation state that EOSC data and EOSC services are intended to be free at the point of use or access[20]. The feasibility of this is related to funding as discussed above. Below we discuss this from a policy point of view, based on the examples of the use cases.

The high-level policy objective of EOSC is to support FAIR data. As far as possible, scientific data should be freely available in the public domain and in reusable formats. However, in order to utilize this data a number of services is frequently required.

In the use cases, services such as SNIC Alvis, SNIC Swestore, and UT Rocket offer services backed by a limited amount of resources. These resources are consumed when used by a researcher. This means that there will be competition among users of the service. If the service is free at the point of use, this competition must be resolved in some other way. On the other hand, data services provided by ICOS ERIC and to some extent CSC IDA offer data that is not consumed when used by researchers. As a result, there is little or no competition. The data services must be funded, and agreements must exist for how this funding is sustained, but the cost and hence the funding needed is not affected by offering the service free at the point of use or by the level of usage.

In the case of ELIXIR, data is tied to a complex research infrastructure offering also resources for utilizing data. This simplifies the work of research groups in the community but can also mean that access is restricted to community members. It can be a challenge to the use of the data from outside the community.

[19] The Wallenberg Foundations is the collective name for the non-profit public and private foundations established by individual members of the Wallenberg family. The foundations support excellence in Swedish research. Annual contribution is in the order of 2 bn SEK. https://www.wallenberg.com/en

[20] Draft EOSC RoP: https://repository.eoscsecretariat.eu/index.php/s/QWd7tZ7xSWJsesn#pdfviewer

EOSC-Nordic project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 857652

5.5 Service Operations and Support

To adequately serve research applications, services require operations capabilities, and researchers must have access to support for the services. For some services, such operations and support can be offered by a single institution. In EOSC, both operations and support is envisaged to be federated, distributed across several actors. This is also the case for some of the national services and services from thematic e-Infrastructures in our service cases.

Distributed and federated operations and support require agreements among the stakeholders, so that roles and responsibilities are clear. They also require commitment to the service from all stakeholders. Such agreements and commitment are in place for some of the national services and for the thematic e-Infrastructures. Similar agreements for federated operation and commitment to the sustainability of the services are possible in EOSC but may require some changes to existing national agreements.

IDA and UT Rocket are both operated and supported centrally, by a single operator. They can be seen as similar in operation to a service from a commercial service provider.

SNIC Swestore is operated and supported in a distributed manner by a number of HPC facilities in Sweden. SNIC Alvis is operated by a single actor but supported by multiple operators, in part so that support can be close to the user communities. A set of national agreements are in place for this, as well as for coordination and oversight. The focus is on ensuring commitment to the services.

ELIXIR is operated and supported in a federated manner by a significant number of national nodes. Each node supports resources that are part of a national contribution to ELIXIR and support national users. In addition, some nodes support resources for all of ELIXIR. Agreements on roles and responsibilities for this federated model exist within the ELIXIR collaboration. ICOS ERIC has a very similar situation with national nodes but it also has one additional layer in its service chain. The Central Facilities are consortia based on several countries, so countries also support common activities to have better data and services.

Service operations and support is a critical aspect of the service delivery chain. It is the motor that ensures that the service is realised in a timely and reliable manner. As can be seen, service operations and support are in some cases simple and centralized. This is typically the case for national services that are restricted in scope. When services become more complicated in scope, and especially as services cross borders, operations and support likewise become complex, and require that agreements on roles and responsibilities are in place. In some existing service cases, such agreements also serve to create commitment to the services within a community.

5.6 Resource Type

The fundamental purpose of services is to offer a resource to the scientific community. These resources are fundamental enablers for the scientific process and are instrumental in allowing research collaboration. Resources differ in how tangible they are, in the level of competition for the resources, and in the duration of the resource.

The resources offered by UT Rocket and SNIC Alvis are tangible (access to a computing), are temporary (last as long as the computer job) and competitive (the resource can support a limited number of users due to scarcity of resources).

The resources offered by SNIC Swestore and IDA are tangible (use of a hardware storage facility), have a significant duration (limited to a research project or program in the case of Swestore, extending past the duration of the research itself in the case of IDA). Both services are competitive as storage is limited by funding and can be used by a large number of researchers.

EOSC-Nordic project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 857652

The resources offered by ELIXIR are both tangible and intangible (computing for data analysis, and the datasets of a science community). Some resources have significant duration for the research datasets; the ELIXIR data is intended to be the permanent repository of a major science community. The data storage and discovery services are therefore long lived. Data analysis services can on the other hand be temporary. Competition for the resources is low, as the resources are dedicated to a specific community and scaled to the needs of that community. There is no competition for storage or compute resources with other science disciplines (though there is certainly competition for funding with other disciplines, a competition that sets the scope of the ELIXIR collaboration).

The resources offered by ICOS are mostly intangible (research data sets). The resources are long-lived (the permanent, consolidated research data sets for atmosphere, ecosystem and ocean data for 12 European countries). The resources competition is low; there is no inherent limit in a data set on how many users or projects can use the data.

The type of resource has implications for the service delivery chain. Intangible and general resources are in general easier to manage in a relation between provider and end user. Other types of resources usually involve other stakeholders. In particular, limited resources with competition require prioritization as part of the service delivery. In a traditional marketplace, the market would handle this prioritization in terms of price. In e-Infrastructures for research, stakeholder engagement and national priorities have a significant role.

5.7 Service governance

Services are governed in a variety of ways. For some services, user engagement in the governance is crucial; for others, funding parties are the main governance entities. In some cases, governance is an element of a series of agreements, maybe in the form of a memorandum of understanding, between stakeholder institutions. In the case of cross-border services, this can have the form of bilateral agreements, or the form of multi-stakeholder agreements. Such agreements - in particular the multi-stakeholder agreements - may have formal and informal components. The informal components, i.e., a community understanding of joint benefit, can be crucial for the sustainability of a set of services.

In the case of UT Rocket and IDA, the services are primarily governed by the institution offering the service, with input from the funding body. Users and operators are stakeholders that may be consulted and who may give input for the evolution of the service but have little formal say in terms of service governance.

In the case of SNIC Swestore, the service is governed by boundary conditions set out in a consortium agreement, grant agreement, third party agreements, allocation guidelines, as a collaboration by committees representing the SNIC member institutions. The governance gives voice to both user communities, hosting institutions as well as funding bodies. The governance and hence the service evolution and sustainability depend on the ability of governance to form consensus among the stakeholders.

In the case of SNIC Alvis, the service is also governed by SNIC governing bodies, governed by boundary conditions set out in a consortium agreement, grant agreement, third party agreements, allocation guidelines, and additional conditions set out by the donor, Wallenberg Foundation.

In ICOS and ELIXIR, the services and the service portfolio are governed as part of the governance for an established, international, multi-stakeholder collaboration. This involves national and European funders, user institutions, and service providers. In ICOS, the governance is embedded in the ERIC structure. In ELIXIR, governance involves the ELIXIR Hub, national heads, node management, as well as a board and a range of governance bodies. In both ICOS and ELIXIR, a crucial part of governance is to facilitate researcher

EOSC-Nordic project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 857652

and research body engagement. This engagement serves to focus the services delivered in a way that maximises benefit for research. A key side effect of the engagement is to increase the commitment of researchers and the research institution to the use and future support of the services.

In all cases, governance serves in part to prioritise resources. This can take the form of deciding who can use limited resources, prioritising evolution of the service, or prioritising the use of stakeholder resources to support the services (and not some other service). In this way, the governance to some extent is a stand-in for the market. By having research communities and individual users engaged in the service delivery chain, resources can be priorities to ensure benefit for research and commitment to the sustainability of the service. Without such engagement, every service is equally important.

5.8 Summary

Services, and their actors exist in the context of traditions, agreements, and commitments between stakeholders - nationally and cross-border - and the human network of the people engaged. Service delivery, looking at the use cases described in this deliverable, has developed over time in institutional contexts that make up ecosystems, based on national actors, or community actors. Service delivery for ELIXIR, ICOS ERIC and similar, large scale, established communities, depends on the engagement of a network of institutional stakeholders. The context of service delivery is particularly important in addressing resource limitation, as it provides the necessary background for prioritization and representation of the sciences using the services.

Some service cases e.g. UT Rocket, IDA, SNIC Alvis, are services provided by national e-infrastructures whose primary mandate is to serve country-specific users. Including these services to EOSC requires negotiations with the governing bodies and changing of agreements. Many services involve consumable resources and hence have competition for use. Presently, such competition is resolved either nationally by national institutions, or within collaborations specific to a research community. It is as yet unresolved how such competition may be resolved for services offered generally through EOSC.

Other services mainly offer access to research data, and do not consume resources. This simplifies the provision of the service through EOSC.

Finally, some services may face additional hurdles beyond those described above. A service may, as in the example of SNIC Alvis, be based on resources made available specifically with a narrow, strategic aim or a national purpose.

EOSC-Nordic project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 857652

6. Conclusions

This report has illustrated, based on six service case studies, that providing ICT services for research requires complex and careful orchestration, both technically and organisationally. This orchestration depends on (1) the interplay between many stakeholders; (2) existing national and international agreements on rights and responsibilities; and, (3) traditions and human networks.

We have considered EOSC from the perspective of the delivery chain. We have studied the relationship between the many stakeholders involved in the delivery of services. We have seen that these relationships are dependent on the nature of the services, the community of users, and the basis for the funding of the services. We have seen also that these relations are especially complex for cross-border service delivery.

We have analysed service case studies based on aspects of the service delivery chain, considering user communities, funding, operations and support, resource types and governance. We have seen how all these aspects of the service delivery chain influence the service offering and its relation to EOSC.

We have shown that the service delivery chain can take many forms; it is different for all of the six service cases considered. We have also seen that all the elements are essential to ensuring that the service is sufficiently funded, suitable for purpose, and sustainable. There is no single, cookie-cutter approach that will make services or their service delivery chain uniform.

The analysis of the six cases suggests that some services are generic and can be offered universally through EOSC, provided that funding can be ensured in order that consumable resources may be replenished. Other services mainly provide access to data, and can be provided through EOSC, without additional funding, if the stakeholders see a benefit in doing so. Some services are national in scope and governance and may be regarded as strategic for national aims. For such services, integration with EOSC will require consultation with national stakeholders. Finally, some services are thematic, specific to a community and integrated within the research practices of that community. Offering these services through EOSC may conflict with existing stakeholder relations in the community and may be seen as offering little benefit to communities with established European collaboration and offering these services through EOSC would therefore require negotiations with the stakeholders.

A successful delivery chain involves many actors (e.g. scientists, funders, service providers, data stewards, policy makers etc) in different contexts (e.g. geographical - local, national, regional, international; organisational - university, research & e-infrastructure, government etc). Each actor plays a critical and/or unique role that is crucial to the smooth delivery of the services. Becoming a part of the EOSC service ecosystem means dealing with additional layer(s) of complexities and interconnections in the delivery chain, which is an aspect that should be planned and executed meticulously. A well-functioning EOSC will benefit European research through increased access to services and reuse of data. Realising this ambition requires that EOSC accommodates a range of service delivery chains, making service onboarding and integration feasible. It is necessary that EOSC is seen to add value to existing service stakeholders and new beneficiaries alike, to encourage engagement for the benefit of European research.

In future work we will, based on the analysis presented here, study the existing cross-border collaboration models in the Nordic and Baltic countries (for publication in February 2021), and with that and the present work in mind propose recommendations for future cross-border collaboration in the context of EOSC (for publication in August 2021).

EOSC-Nordic project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 857652

List of Abbreviations

Consortium agreement (CA)
IT Center for Science (CSC)
Danish Technical University (DTU)
Danish e-Infrastructure Cooperation (DeIC)
Estonian Scientific Computing Infrastructure (ETAIS)
European Open Science Cloud (EOSC)
European Strategy Forum on Research Infrastructures (ESFRI)
European Research Infrastructure Consortium (ERIC)
Findable, Accessible, Interoperable, Reusable (FAIR)
High-performance computing (HPC)
Grant Agreement (GA)
Higher Educational Institutions, (HEI)
Integrated Carbon Observation System (ICOS)
Information and Communication Technology (ICT)
Latvian Ministry of Education and Science (IZM)
Latvian academic network (LAT)
National research and education networks (NREN)
Principle investigator (PI)
Rules of Participation (RoP)
Swedish National Data Service (SND)
Swedish University computer Network (SUNET)
Swedish National Infrastructure for computing (SNIC)
University of Tartu (UT)

EOSC-Nordic project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 857652

Appendix A: Actors, Roles and Responsibilities

ELIXIR

Example scenario

A Finnish researcher wants to access and use data for Protein modelling. One option is UnitPro, a comprehensive resource for protein sequence and annotation data[21]. The researcher logs into ELIXIR bio.tools platform and gets access to the needed resource.

Stakeholders, roles and responsibilities of the delivery chain in this scenario

Stakeholder	Role(s)	Responsibilities
European Bioinformatics Institute (EMBL-EBI),	Service providers	Collaborate Perform tasks such as database curation, software development and support.
SIB Swiss Institute of Bioinformatics		
Protein Information Resource (PIR).		
ELIXIR Node	Node Coordinator	*Apply for ELIXIR Core Resource status
ELIXIR hub	ELIXIR Coordinator	*Assess the application
Researcher	Service/Data consumer	Access and use the resource through the bio.tools platform
Danish ELIXIR node	Developer Long-term supporter	Lead the development of bio.tools Ensure long-term support in collaboration with other European experts as well as ensure that *bio.tools* remains free, open and maintained in the long term.

*This happens once when there is an application for inclusion to the ELIXIR Core Data Resources list.

[21] ELIXIR Core Data Resources: https://elixir-europe.org/platforms/data/core-data-resources

EOSC-Nordic project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 857652

UT ROCKET

Stakeholder	Role(s)	Responsibilities
UT	Service owner	Service owner is responsible for correct and accurate data (service description) of the service
UTHPC	Service provider	Manages and delivers the service/resource database curation, software development and support.
		Act in accordance with commonly agreed principles regarding the conduct of research
		Should not willfully misrepresent or provide false data
		Publish the terms of use for their services
		Define and publish the quality targets of their services
ETAIS	Authentication service provider	Provide authentication service
EOSC marketplace		Provide authentication service
	Service consumer	Access and use the resource through the EOSC marketplace or directly from ETAIS self service portal https://minu.etais.ee using eduGain, TaaT or Riigi Autentimisteenus authentication

IDA

Stakeholder	Role(s)	Responsibilities
The Ministry of Education and Culture (Finland)	Service owner Funder	Avail funds for service development and support Monitor progress
CSC – IT Center for Science Ltd.	Service producer/provider	Provide the service to users at the agreed service level. Provide support in the agreed manner. Notify user organisations and users of all security breaches and security risks without delays.

EOSC-Nordic project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 857652

Project (group of users belonging to the same shared storage space)	Support group membership	Access and use the storage space
User	A project member	Obtain a CSC account and project membership
Project manager	Take charge of the a specific project	Manage group's access, data and service usage Primary communication contact person with CSC.

ICOS ERIC

Stakeholder	Role(s)	Responsibilities
National Providers		
Organisations in different European Countries owning the stations.	Data providers	Provide data according to the rules and licenses decided by the ICOS ERIC.
Central Facilities		
CEA, CNRS, USVQ as the French partners and Finnish Meteorological Institute as the Finnish partner of the ATC. Fondazione CMCC in Italy, University of Antwerp in Belgium and INRA in France as partners of the ETC. University of Heidelberg and Max Planck Institute in Germany as partners of the CAL. University of Exeter, United Kingdom Research & Innovation through National Oceanography Centre of Waterfront Campus, Plymouth Marine Laboratory as the British partners & Norwegian Research Centre & the University of Bergen as the Norwegian partners of the OTC	Service providers for the RI partners	Collaborate Process and produce elaborated ICOS data, including e.g. data quality control and gap filling and provide other agreed services to National Networks according to the rules and licenses decided by the ICOS ERIC.
ICOS ERIC		

EOSC-Nordic project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 857652

Lund University and Wageningen University in the Netherlands as the partners providing the services of the ICOS ERIC Carbon Portal. Member countries of ICOS ERIC.	ICOS ERIC coordinates all activities and the Carbon Portal part of ICOS ERIC provides access to data, data tools and metadata for users. Member countries provide funding to maintain ICOS ERIC and also the services of the RI.	ICOS ERIC oversees that the data and data tools provided by the National Networks and the Central Facilities are provided according to the agreements and decided rules and licenses of ICOS ERIC. Member countries make decisions on what conditions data is available from ICOS RI and ensure sustainable funding for all services.

SNIC Alvis

Example scenario

A researcher would like to use the SNIC resource for research using AI/ML techniques. The PI, if eligible, can apply for the resource via the SNIC User and Project Repository. The application is reviewed by a reviewer at any of the SNIC members, e.g. Lunds University.

Stakeholders, roles and responsibilities of the delivery chain in this scenario

Stakeholder	Role(s)	Responsibilities
Swedish Research Council	Funder	Financing
KaW		
Uppsala university	SNIC	Coordination. Publish the terms of use for their services
Uppsala university, Umeå university, Chalmers, KaW/WASP	Governing board	Making decision
Chalmers	Service producer/provider	Provide the service to users at the agreed service level. Notify user organisations and users of all security breaches and security risks without delays.
Umeå university		
Uppsala university		
Researcher	Service/Data consumer	Obtain a SUPR account and

EOSC-Nordic project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 857652

		project membership. Access and use the resource.
Umeå university, Uppsala university, KTH, Linköping university, Chalmers, Lund universtity	Allocation and user support	Enablement
Chalmers, Linköping university	Monitoring and accounting	Monitor and report on the level of usage of data / services available.

SNIC Swestore

Example scenario

A researcher would like to store data, perhaps to be used in different locations. A researcher would like to use the SNIC resource for research using AI/ML techniques. The PI, if eligible, can apply for the resource via the SNIC User and Project Repository. The application is reviewed by a reviewer at any of the SNIC members participating in the Swestore dCache operations.

Stakeholders, roles and responsibilities of the delivery chain in this scenario

Stakeholder	**Role(s)**	**Responsibilities**
Swedish Research Council	Funder	Financing
Uppsala university	SNIC	Coordination. Publish the terms of use for their services
SNIC Board	Governing board	Making decision
Chalmers	Service producer/provider. Own the RoP.	Provide the service to users at the agreed service level. Notify user organisations and users of all security breaches and security risks without delays.
Umeå university		
Lund university		
Linköping university		
Researcher	Service/Data consumer	Obtain a SUPR account and project membership. Access and use the resource.

EOSC-Nordic project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 857652

Umeå university, Uppsala university, KTH, Linköping university, Chalmers, Lund universtity	Allocation and user support	Enablement
Chalmers, Linköping university	Monitoring and accounting	Monitor and report on the level of usage of data / services available.

EOSC-Nordic project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 857652

Appendix B: Case Descriptions

CASE: ELIXIR Bioinformatic RI

The case and relation to EOSC: ELIXIR is the most important biological information infrastructure in Europe as well as one of the most important infrastructures globally, connecting scientists across the globe with a vast amount of bioinformatics data. The data is freely accessible to anyone. The usage is based on a mutual understanding between the scientific community, to adhere to similar data structures - including metadata standards, ontologies and vocabularies - and the willingness to upload one's own data to a common data platform. In return the scientists help the continued accumulation of data that can be utilized and referenced freely.

ELIXIR RI is a relatively new organization, founded in 2013. It rests heavily on work performed by the European Molecular Biology Laboratory (EMBL) and its outstations The European Bioinformatics Institute, UK as well as the Swiss Bioinformatic Institute, CH that for many years stored the bioinformatic data produced in Europe and globally on a voluntary basis. Since technology has improved, specifically the speed and cost effectiveness of DNA sequencing technology, so have also the complexity of coordinating efforts and the cost of hosting data. For that reason, efforts were made to put a new Bioinformatics Research infrastructure on to the first ESFRI list. This work was completed by 2013 via establishing ELIXIR RI as an International Treaty Organization, made up of around 30 countries and EMBL, all participating in the funding and governance of the infrastructure.

Funding

The ELIXIR Hub is funded through membership fees paid by Member countries, and majority of this funding is transferred back to the Nodes to implement ELIXIR's five-year Scientific Programme. In 2017 this amounted to around 5 MEUR in total. A substantial part of the funding comes from competitive grants, mainly from H2020 and IMI. ELIXIR Nodes, which run the services, are typically funded through national-level investments, either through dedicated infrastructure grants or through competitive research grants. Private foundations also support ELIXIR Nodes. The total cost of operating ELIXIR has been estimated to be well in excess of 100 million euros, involved in running all of the 140+ ELIXIR services, across 220+ legal entities that form the ELIXIR Nodes.

How ELIXIR works

ELIXIR is an intergovernmental organisation that brings together life science resources from across Europe. These resources include databases, software tools, training materials, cloud storage and supercomputers. The goal of ELIXIR is to coordinate these resources to form a single infrastructure, providing access to data and IT resources, as well as driving standards and best practises.

Resources

Data platform

ELIXIR coordinates access to trusted data resources. The ELIXIR Data platform promotes Open Access and has terms of use or a licence that enables the reuse and remixing of data within a robust, long-term sustainable, coordinated, scalable and connected data ecosystem. The data platform consists of 19 core data resources, and additional 80 thematic databases and data deposition resources coordinated by ELIXIR and maintained locally with ELIXIR partners. ELIXIR provides a registry of data resources through fairsharing.org run by the University of Oxford.

Tools:

EOSC-Nordic project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 857652

Analyzing vast amounts of bioinformatics data requires specific software tools for high-throughput computation. Access to these tools (i.e. algorithms) are equally important as the access to the data in order to both reuse the data and to validate research results. The tools are openly available through the ELIXIR partners. Since there are several thousands of these available the Danish ELIXIR node has the role of providing a searchable and linked registry of tools.

Compute platform

The ELIXIR Compute Platform was established to build and integrate cloud, compute, storage and access services for the life-science research community. The data needs to be managed as a federation, where data providers work as a single infrastructure. This enables providing mechanisms for researchers to move their analysis to where the data is located from different sources across Europe. The objective is to combine all the components of the ELIXIR Compute services into a seamless workflow. ELIXIR works with services provided by the ELIXIR Nodes[22] and their partners, and by the European e-Infrastructures (i.e. EGI, EUDAT, GÉANT, EOSC, Helix Nebula Science Cloud.

Interoperability and Standards

The Interoperability Platform aims at supporting the life science community in adopting standardised file formats, metadata, vocabularies and identifiers. This maximises the value and benefit by integrating data from disparate resources across disciplines and borders by developing the FAIR service infrastructure that aligns different aspects of tools, metadata, ontologies, standards, workflows and identifiers that are fit-for-purpose.

Training

ELIXIR coordinates training activities and resources across the ELIXIR community and runs a registry in which trainers and students can upload and find resources and training courses available. Also an online e-learning portal is available. Thus ELIXIR provides tool kits, training materials and a technical platform for learning within the community.

Governance:

ELIXIR Board: the main decision-making body in ELIXIR, consisting of members appointed by the member states.

Scientific Advisory Board: advises the Board on ELIXIR's scientific strategy and reviews the ELIXIR node applications.

Industry Advisory Committee: gives advice and guidance on industry needs.

Director: responsible to the ELIXIR Board for implementing the ELIXIR's scientific programme.

ELIXIR HUB: the ELIXIR main office, responsible for coordinating efforts across the ELIXIR organization.

Heads of Nodes Committee: consists of the Director and the Heads of the ELIXIR national infrastructures (Nodes) - responsible for developing the ELIXIR's scientific and Technical strategy and scientific programme.

ELIXIR Nodes: nationally organized entities responsible for carrying out specific tasks for ELIXIR and for national organizing of the community,

Communities: ELIXIR organizes communities, in order to bring scientists within specific fields or domains together and support build up of services and computational resources. Currently there are twelve communities, e.g. proteomics, plant biology, federated human data and rare diseases,

Commissioned services: commissioned services are a tool where ELIXIR sets up a task force, and allocates resources to address key issues of importance for the continuing development of the RI. Thematically this

[22] https://elixir-europe.org/about-us/who-we-are

EOSC-Nordic project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 857652

currently involves the following issues: literature-data integration, long term sustainability, biocuration - practice capability and training, scaleable curation, support for core data and deposition resources.

Summary

ELIXIR RI is a truly distributed and federated data infrastructure, with many actors representing international, national and local institutions, having many different roles and responsibilities in the service delivery chain. There are literally hundreds of entities and thousands of people involved in the service delivery in a managed and controlled manner.

The key driver is the challenge of deriving knowledge from huge amounts of biological information - a challenge that requires storage and computational power, as well as community acknowledged data standards ontologies and metadata which enables this co-operation. The data challenge has pushed a very diverse community into establishing a digital infrastructure and utilize services that no single institution or even a nation can manage to deliver on their own.

The challenge, however, is that this level of integration requires an enormous coordination effort and some very careful prioritization in terms of services, balancing the need to cater to "core users" and big communities (eg. Human biologists) as well as catering to smaller communities and users with basic knowledge of bioinformatics. Also, ELIXIR is a "big ship" to turn, in terms of culture, technology and data regimes in trying to shift focus towards a truly federated data resource based on machine readability / interoperability.

CASE: UT Rocket

The case and relation to EOSC

UT Rocket is an Estonian service delivered by the University of Tartu. UT Rocket is primarily an HPC batch processing cluster. Most of the resources are provided using slurm batch queue, bioinformatics workflows are also available through local Galaxy portal. It is also possible to use GPU-s and singularity containers. The cluster is owned by the University of Tartu. The cluster operating costs are covered from usage fees. As a practical matter it means that while the cluster is accessible to everybody who wishes to perform scientific work (primarily, but not limited to Estonian R&D institutions, also commercial and public institutions) all users must belong to some organization that covers their usage costs. UTRocket supports local, EduGain and EU eID authentication.

In relation to EOSC, it is possible to provide free of charge usage at the point of access, provided such access is financed by some program (pilot projects, virtual access (VA)).

How UT Rocket works / Example Scenario

A user from Taltech wants to use the UT Rocket cluster for Baltic Sea simulations. The User requests access through EOSC marketplace or directly from ETAIS self-service portal.[23] The owner of the Taltech organization grants access to the user on ETAIS portal. The user can create an allocation on the rocket cluster and create her FreeIPA account. The user can now access the rocket cluster. She can use existing software on the cluster, install her own or request the software to be installed by support staff. The user can then perform her analyses and use mid-term storage at UTHPC or transfer the results back. The usage statistics and accounting are available on the self-service portal and can be reported externally through an API when necessary. Upon publishing the results, it is possible to reference UTROCKET by its doi number.

Resources

[23] https://minu.etais.ee

EOSC-Nordic project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 857652

The main part of the UT Rocket cluster consists of 135 compute nodes, two compute nodes with GPUs and a head node. In addition to these nodes, there is a large memory machine with 2TB of RAM and two GPFS filesystem servers, which will provide fast storage for the entire cluster. All the machines mentioned above are connected to a fast Infiniband 4X QDR fabric.

Governance

The use of UT Rocket cluster is mainly governed on the following levels:

a) The ETAIS consortium agreement, which declares, that each resource owner/ETAIS partner can enforce their own access and usage terms. The ETAIS consortium is governed by the ETAIS board, which consists of ETAIS Director (*ex officio*), ministry representative, and representatives of partner universities and NREN.

b) As the UT Rocket cluster belongs to UT, its usage is subject to UT HPC consortium statute and price list. UT HPC consortium is governed by representatives from various UT departments who represent the majority of local users.

c) Finally, as most of the hardware has been procured using EU funding, it is subject to rules governing state aid and EU level restrictions (meaning that no more than 20% of ETAIS consortium resources can be made available commercially).

Summary

UTROCKET is an example of a classical HPC service. Its small size might make it impractical for "real" work, at least for bigger countries, but ideal for all kinds of pilots. It's rules of operation have relatively few restrictions and decisions can be made on an executive management level, supporting high speed development. It is also already widely integrated into all kinds of federations (EduGAIN, EuID). It will also evolve to give as close user experience to the EuroHPC LUMI supercomputer as possible.

CASE: IDA

The case and relation to EOSC

IDA is a component of the Fairdata family of services providing integrated solutions for storing, sharing and publishing research data. IDA specifically enables continuous and safe storage of research data. Data can be saved, organised and shared within a project group and further stored in an immutable state. Hence, IDA is intended for files that are not used and modified on a daily basis. IDA is currently not listed in the EOSC catalogue. However, it has the potential to fulfill the demand for storage services from a wider European user base under the conditions stipulated by CSC for researchers working from abroad[24].

How IDA works and example scenario

The storage space in IDA can be shared among the project members, enabling e.g. collaborative projects to utilise a shared storage space not tied to a single organisation.

A principal investigator (PI) in a Finnish university is in need of storage space for a project. He/she applies for an IDA storage quota, and usually becomes the project manager in charge of the storage space. The IDA contact person in the organisation selected by the project manager processes the application. Once the application is approved, the project manager can access the space, and has the right to add and remove users. A user first needs to register for a CSC user account, after which the project manager can grant access to the storage space. A project can have members from multiple organizations and also collaborators from abroad.

Resources

[24] Other registration options: https://research.csc.fi/accounts-for-researchers-working-abroad

EOSC-Nordic project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 857652

CSC produces and hosts the IDA storage capabilities on CSC's servers. IDA users can use the storage space from both their own computers and from the servers hosted by CSC. The storage quotas range from 1 GB to 100 TB. The storage of files in IDA can be managed using the web and command line client interfaces; however, the contents of the stored files can't be modified directly. Instead, a stored file must be first retrieved from IDA to either CSC supercomputers or some other computer in order to analyse or modify the data.

Governance

IDA is funded by the Finnish Ministry of Education and Culture and produced by CSC – IT Center for Science Ltd. It is dedicated for research purposes and offered free of charge to users and researchers in all higher education institutions and state research institutes in Finland. Each organisation has an IDA contact person who is responsible for reviewing and approving/rejecting IDA storage applications from users in each respective institution.

Summary

IDA enables organising, sharing and storing stable research data. The data is typically represented in the form of datasets. A published research dataset gets a persistent identifier (DOI or URN) and a landing page, making it findable and reusable. The actual data (files) in the published dataset can be set as openly downloadable by others, but it's also possible to publish only the metadata of the dataset. By its very nature, IDA enables the F and R of FAIR, hence upholding EOSC's vision[25] to increase findability and reusability of datasets.

SNIC services

The below section details two use cases from the swedish national infrastructure for computing (SNIC), aiming to illustrate service delivery of a distributed service, and a privately funded resource.

CASE: SNIC Swestore dCache

The case and relation to EOSC

Swestore dCache is a distributed storage service aimed at active non-sensitive data with low to moderate performance needs.

How Swestore dCache works and example scenario

A researcher would like to store data, perhaps to be used in different locations. A researcher would like to use the SNIC resource for research using AI/ML techniques. The PI, if eligible, can apply for the resource via the SNIC User and Project Repository. The application is reviewed by a reviewer at any of the SNIC members participating in the Swestore dCache operations.

Resources

There are mirrored metadata servers at two sites (Umeå University and Linköpings University). Data storage is available at four sites (Umeå, Linköping, Chalmers University of Technology, and Lunds University). Data can be stored/retrieved using customary dCache methods, with or without certificates. Data is always stored in two different geographical locations. The service is not intended for sensitive data.

Governance

[25] https://ec.europa.eu/research/openscience/pdf/eosc-fair_paper_schouppe-burgelman_2018.pdf

EOSC-Nordic project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 857652

Swestore dCache is a SNIC Service with a budget decided by the SNIC Board. There is an activity leader for storage within SNIC, who has help from two coordinators specialising in dCache.

Summary

Swestore dCache is a distributed service, with resources owned by four different legal entities, governed by third party agreements set up between Uppsala University as the host for SNIC and the consortium parties hosting the service. The distributed ownership means that data is spread across four legal entities, each responsible for legal matters and the level of security that must be upheld for the resource. They all have their own procurement processes, and reinvestments should be aligned with other investments at the site. In order to resolve support issues there is often a need to look into the meta-data servers, which requires system administration privileges. The division of labor between the parties is fixed by third party agreements.

Should Swestore dCache be added as an EOSC service, a decision has to be worked out with the Research Council, as the grant agreement for SNIC stipulates that a PI must be affiliated with a Swedish research organisation. A decision by the SNIC board to allow for this additional use would also have to come in place.

CASE: SNIC Alvis

The case and relation to EOSC

Alvis is a compute resource for AI/ML research. There is a repository of data sets relevant for AI research available by the resource. Alvis is funded by the Knut and Alice Wallenberg Foundation through their WASP program. Operations are funded by SNIC.

How Alvis works and example scenario

A researcher would like to use the SNIC resource for research using AI/ML techniques. The PI, if eligible, can apply for the resource via the SNIC User and Project Repository. The application is reviewed by a reviewer at any of the SNIC members, e.g. Lunds University.

Resources

Alvis is set up as a compute resource with a number of partitions based on different accelerators, currently GPU:s. It is connected to a CEPH-based storage system shared with a traditional HPC system and the SNIC Science Cloud, providing file system storage as well as object storage. The resource is owned by Chalmers University of Technology and system administration is shared between parties within the SNIC consortium. User support is provided by other parties of the SNIC consortium as well, e.g. The Royal Institute of Technology.

Governance

There is a steering committee consisting of the SNIC director, a WASP representative, a representative from Chalmers, and one other person. The placement of the resource was decided by the SNIC Board. There is a reference group, representing researchers. There is also a group, consisting of one person connected with WASP, one working with allocations, and one close to operations assisting with specifications for procurements. Chalmers has provided a project leader who is attending daily matters.

Summary

The resource is owned by one legal entity and governed by third party agreements set up between Uppsala University as the host for SNIC and the consortium party hosting the service.

To allow service provisioning outside the scope of the current user base conditions set by stakeholders, both inside and outside the SNIC consortium, would need change. System administration is done by three different parties, all of whom must comply with the requirements of the system owner. Allocation of the

EOSC-Nordic project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 857652

resource must not only fulfill the requirements of the donor, KAW, but also of SNIC and thus the Swedish Research Council. Allocation is handled by an allocations committee, which initially is distributed across six Swedish universities.

Should Alvis be added as an EOSC service, several agreements have to be altered. First and foremost, the conditions of the donation are presently aimed at Swedish research. The grant agreement between the Swedish Research Council and SNIC mandates a PI affiliated with a Swedish research organisation. Thus, an understanding has to be worked out with the Research Council. This presupposes a SNIC Board decision. There are also procedures regarding allocation that have to be modified.

CASE: ICOS ERIC

The case and relations to EOSC

The Integrated Carbon Observation System, ICOS operates in three domains, namely atmosphere, ecosystem and ocean[26]. It consolidates open data from more than 130 measurement stations across 12 European countries. The stations observe greenhouse gas and carbon activities across the atmosphere, land surface and the oceans.

The ICOS community is composed of a wide network of scientists and professionals from different universities or institutes in member and observer countries. The community collaborates with external partners and helps fight climate change and address its impacts in a sustainable manner.

ICOS is part of the EOSC-hub project, where it has a dedicated task to establish a competence center on the management of measurement station information together with the European Long-Term Ecosystem (eLTER) Research Infrastructure[27].

How ICOS works and example scenario

ICOS manages a Data Portal through which users can access the data. All the data is assigned a landing page containing information about the data. The data is made available for downloading after it has been assessed and checked for quality.

A scientist wishes to download a dataset on greenhouse gases. It's possible to download the data from the ICOS data portal after accepting the Data Policy and Data Licencing Agreement (without having to register). This is convenient for a one-off instance. However, should a user need to download data regularly, it is better to create a user account. The user can also visualize the data, if needed. When logged in, a user can save his/her datasets in a shopping cart, and can retrieve the data later.

Resources

The **ICOS Data Portal** enables users to download and visualise data on greenhouse gases. There are a number of additional tools to further support data visualisation, management and analysis. These are:[28]

> *STILT Footprint:* An online tool to analyse the potential impact of natural and human-caused emissions on the atmospheric carbon dioxide at a selection of ICOS atmosphere stations.
>
> *Jupyter notebook*: A collection of tools for interactive computing and sharing of computational ideas in a virtual research environment.

[26] ICOS in a nutshell: https://www.icos-cp.eu/about/icos-in-nutshell

[27] Projects: https://www.icos-cp.eu/observations/projects

[28] Tools: https://www.icos-cp.eu/data-services/tools

EOSC-Nordic project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 857652

DOI minting: Supports attaching Digital Object Identifiers (DOIs) to the data.

SPARQL endpoint: Enables access to the metadata.

Statistics for data downloads: Browse Data Portal's download statistics.

Upload data portal: Allows requests to perform manual data uploads to the ICOS Data Portal.

Governance

ICOS ERIC's member countries finance the ICOS ERIC activities and hence decide what services can be available. The member countries also ensure that the national services like Central facility and station services are funded.

National networks are stations in different countries providing raw data. In order to become an ICOS station, the stations need to go through a labelling process.

The ICOS ERIC Head Office coordinates the RI activities and the labelling process of the stations. ICOS ERIC's Carbon Portal provides access to ICOS labelled data and data tools as well as metadata.

Central Facilities are consortia which process data further in order to have quality ICOS data and data tools. These Central Facilities are, Atmosphere Thematic Centre (ATC), Ecosystem Thematic Centre (ETC), Ocean Thematic Centre (OTC) and Central Analytical Laboratories (CAL).

Summary

ICOS is an RI with a mission to combat climate change and its impacts. It provides high-quality, open greenhouse data and services across Europe. ICOS links research, education and innovation ecosystems in efforts to further scientific excellence. Since 2015, ICOS has been in operation as a recognised European Research Infrastructure Consortium (ERIC). As such, ICOS ERIC is a distributed research infrastructure, similar to ELIXIR RI. It's delivery chain involves many actors including scientists (researchers), member countries, stations, universities and institutions. The services are well coordinated and supported by a financial plan to ensure sustainable operations.

ICOS ERIC has established a respected identity that is valued by the community. For example, its biennial conference demonstrates a commitment not only to manage distributed actors, but also to bring them together for knowledge sharing and exchange.

In the context of EOSC, ICOS ERIC has been involved in the EOSC-hub project. Promising results in a pilot project led to an agreement between the EUDAT CDI and ICOS for continued collaboration even after the end of the EOSC-hub project[29]. EUDAT CDI will provide storage and computing services for ICOS.

[29] https://www.eosc-hub.eu/news/eudat-cloud-and-data-preservation-solutions-support-icos-mission

EOSC-Nordic project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 857652