EOSC NORDIC

D2.6: Cross-border Collaboration in the Context of EOSC

Author(s)	René Belsø, Per-Olov Hammargren, Lars-Owe Ivarsson, Pirjo Kontkanen, Leah Riungu-Kalliosaari, Lars Fischer, Sara Garavelli
Status	
Version	Final
Date	4.02.2022

Document identifier:	
Deliverable lead	NORDUnet
Related work package	WP2
Author(s)	René Belsø, Per-Olov Hammargren, Lars-Owe Ivarsson, Pirjo Kontkanen, Leah Riungu-Kalliosaari, Lars Fischer, Sara Garavelli
Contributor(s)	
Due date	31.12.2021
Actual submission date	4.02.2022
Reviewed by	Damien Lecarpentier, Jan Meijer
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	

Executive summary

A key element of open science, FAIR, and the EOSC vision is the notion of cross-border resource access, i.e., use of cross-border infrastructure and services. For EOSC to be successful, cross-border utilisation, provision and usage of resources must be sustainably funded, in a way comparable to what is organised nationally in most Member States. This leads to the discussion about cross-border funding models, cross-border cost recovery, or simply collaboration models.

Different types of cross-border collaborations exist, organised in different ways, and sometimes providing various types of resources to different user communities they need to serve. In the Nordics there are several examples of cross-border collaborations: DICE, ELIXIR, LUMI, ESS, ICOS ERIC, NORDUnet. The EOSC Nordic demonstrators are also examples of cross-border resource provisioning.

These cross-border collaborations cannot be directly compared to EOSC because they have different scopes and they are dealing in many cases with the cross-border provision of one specific type of resources, but many lessons can be drawn and become useful recommendations for EOSC such as the following ones:

- Trying to identify a single funding model for EOSC is very challenging and not an appropriate solution. Breaking EOSC down into different (more clearly defined) components that can be aggregated and provided internationally, according to different strategic rationales and different funding schemes would be the best strategy to move forward;
- Identifying more clearly what is in the scope of EOSC (type of resources and value proposition of offering) can facilitate the identification of appropriate funding models;
- EC funding for EOSC should be substantially complemented by national funding linked to the specific investments, preferably articulated in connection to national strategic interests;
- A working cross-border funding model for generic infrastructures must be found. The LUMI and NORDUnet funding models can serve as starting points;
- Federation and existing legacy should not prevent the selection of the best qualitative and cost-effective solution or to build a new one;
- Consider the LUMI funding model to solve the EOSC computing infrastructure funding issue;
- Some share of funding in EOSC should be dedicated to the cross-border infrastructural needs of the long tail of science;
- Any new funding mechanisms identified by EOSC should come with clear rules and guidelines;
- The new governance framework for digital transition launched in March 2021 by the European Commission for funding multi-country projects deserves investigation by EOSC.

The outputs of this analysis will be shared with the Financial Sustainability Task Force of the EOSC Association as a contribution to the discussion on funding models for EOSC.

Table of contents

Table of Contents

1. Purpose and scope of the deliverable

This deliverable aims to give an overview of the different cross-border collaboration models in place in the Nordic countries for the provisioning of services and resources to support researchers, analysing their funding models and their resource allocation structures from different perspectives: procurers vs users vs providers vs funders. The goal is to provide usable recommendations and input for the future developments of EOSC.

In particular, the recommendations for EOSC will revolve around best practices from the Nordics that could be replicated or piloted at the European level to speed up the establishment of an EOSC funding model and lessons learnt from the Nordics that can be brought to the European discussion and analysis of the existing challenges that deserve attention not only in the Nordic region.

Originally this deliverable was supposed to review and assess the value and impact of the collaboration model proposed by EOSC in the Nordic and Baltic countries and assess the costs of opening up services in EOSC[1], but at the time of writing the deliverable, an EOSC model has yet to be defined and established.

Thus, to formulate the revised scope of the deliverable, two main pieces of advice have been taken into account:

1. the recommendation received in the EC interim review report: "*Furthermore, to foster cross-border research, the project should also provide recommendations for cross-border funding programs*. *Such recommendations are lacking from D2.2, which analyses issues associated with cross-border research initiatives in the targeted region*";
2. the advice received by the chairs of the newly established EOSC Association Task Force on "Financial Sustainability"[2]. The objective of the Financial Sustainability Task Force is to produce a proposal for long term financial sustainability of the main building blocks of EOSC: EOSC-Core, EOSC-Exchange and the Federation of Data & Data Services as defined in the FAIR Lady report "Solutions for a Sustainable EOSC"[3]. The EOSC Nordic project reached out to the chairs of the Task Force to ensure that the outcomes of the deliverable are usable to foster progress in the work of the Task Force. The chairs welcomed the EOSC Nordic proposal and endorsed the idea of focusing the deliverable on analysis of funding models behind existing cross-border collaborations in the Nordics. EOSC-Nordic will submit the deliverable as input to be considered by the Financial Sustainability Task Force.

In addition to analysing the funding models of established cross-border collaborations, the deliverable also reports on the perspective of smaller and less structured research groups in need of cross-border infrastructure services and resources leveraging the experience and the work performed by the EOSC Nordic WP5 demonstrators.

[1] Please note that lessons learnt and reflections on EOSC service onboarding have been analysed in D3.5.
[2] https://eosc.eu/sustaining-eosc
[3] https://op.europa.eu/en/publication-detail/-/publication/581d82a4-2ed6-11eb-b27b-01aa75ed71a1

Despite their recognised relevance for the research sector, the deliverable does not address commercial services. This choice was taken to avoid overlaps with ongoing initiatives such as the OCRE[4] and EOSC Future[5] projects that are specifically addressing this topic.

2. Methodology

This deliverable has been produced according to the following methodology:

- Literature review of relevant documents related to EOSC funding and business models. In particular:
  - Solutions for a sustainable EOSC - A FAIR Lady (olim Iron Lady) report from the EOSC Sustainability Working Group[6]
  - EOSC-hub D2.5 Final Governance and Sustainability implementation roadmap[7]
  - EOSC-hub cross-border service briefing paper[8]
  - EOSC-hub D12.3 Business models and procurement: evaluation and recommendations[9]
  - EOSC glossary[10]
- Analysis of the outcomes of the previous EOSC Nordic related deliverables:
  - D2.2: Cross-Border Collaboration Models – The Nordic Experience[11]
  - D2.4: The EOSC Delivery Chain[12]
- Interviews with representatives of the EOSC-Nordic demonstrators (See Annex 3). The EOSC-Nordic demonstrators involve research groups working in a variety of disciplines, from climate/modelling data to medical data for personalised medicine. Despite the variety of data types, the common denominator among the use cases is the ambition to increase the scientific potential and impact of their research by sharing data over wider geopolitical regions and domains, possibly across disciplines and communities. The interviews were targeted to understand:
  - how the technical components used in the demonstrators were financed for the use case and how they are usually provisioned to other users (*e.g. free of charge, pay per use, etc.)*;
  - if the technical components are already used across borders and if not, the perceived costs of opening them up to EOSC;
  - their views on what could be the most suitable funding models for procuring and provisioning such components.
    The interview template is available in Annex 2.

[4] https://www.ocre-project.eu/
[5] https://eoscfuture.eu/
[6] https://op.europa.eu/en/publication-detail/-/publication/581d82a4-2ed6-11eb-b27b-01aa75ed71a1
[7] https://www.eosc-hub.eu/deliverable/d25-final-governance-and-sustainability-implementation-roadmap
[8] https://www.eosc-hub.eu/sites/default/files/EOSC-hub%20Briefing%20Paper%20-%20Provision%20of%20Cross-Border%20Services%20-%20final_0.pdf
[9] https://www.eosc-hub.eu/deliverable/d123-business-models-and-procurement-evaluation-and-recommendations
[10] https://zenodo.org/record/4472643#.YNQztm5RVEI
[11] https://eosc-nordic.eu/kh-material/d2-2-cross-border-collaboration-models-the-nordic-experience/
[12] https://eosc-nordic.eu/kh-material/deliverable-2-4-the-eosc-delivery-chain/

- Analysis of the different funding models for the existing cross-border collaborations in the Nordics leveraging on the knowledge of the EOSC Nordic partners and filling the gaps by reaching out to external experts directly involved in the collaboration.

The information collected has been analysed by the EOSC-Nordic WP2 team, using the diverse and substantial expertise the group brings. The chairs of the EOSC Association Financial Sustainability Task Force have also been consulted in the scoping phase of the deliverable.

3. Established cross-border collaborations in the Nordics

3.1. Types of collaborations

A key element of open science, FAIR, and the EOSC vision is the notion of cross-border resource access, i.e., use of cross-border infrastructure and services. For EOSC to be successful, cross-border provision and usage of resources must be sustainably funded, in a way comparable to what is organised nationally in most Member States. This leads to the discussion about business models, cross-border funding models, cross-border cost recovery, or simply collaboration models.

However, different types of cross-border collaborations exist, organised in different ways, and sometimes tailored to the different types of resources they need to organise or the different user communities they need to serve. Finding a business model that is appropriate for EOSC as a whole (which deals with different types of resources, providers and user communities) might not be as straightforward as it would seem in the first place. Instead, it is useful to consider various business or funding models for different categories of collaborations. For example, a specific funding model following a pay per use approach may be optimal for brokering the delivery of Infrastructure-as-a-Service from a commercial cloud vendor to researchers in multiple countries, while a model based on national financing and cost-recovery may be suitable to support the activities of a research group based in three countries, relying on a shared data repository hosted in one country, but co-financed by all participating countries.

Furthermore, the landscape of e-Infrastructures is made up of generic and subject-specific resources. Generic resources are often national resources such as networks, national storage or compute resources, but can also be funded by the EC, or by multi-country consortia. Generic resources will often have a policy and process for granting access, such as a scientific merit evaluation for national HPC resources. Subject-specific resources are resources targeted at a specific science domain. These may be national or institutional, available to specific research groups, but can also be European or consortia funded, allowing cross-border research collaborations in the specific domain to utilise shared resources.

To bring some clarity into such discussions, the following matrix will propose categories of cross-border resource- or service collaborations, allowing analysis to proceed for each category, or specific categories to be addressed to find optimal or recommended models. It is hoped that

a development of such a categorisation will allow a more focused discussion of cross-border business models for various types of resources and collaboration models.

To break down the plethora of cross-border collaboration schemes existing in the Nordics into categories, the matrix is organised according to:

- Resource ownership: Federated, jointly owned, owed by the EC
- Resource type: network, storage, compute, data, software, instrumen

With these two categories, we can form a 2-dimensional table:

	Domain specific		**Generic e-Infrastructure**	
	Nationally funded, federated	Nationally funded, jointly owned	Nationally funded, jointly owned	EU funded
Network			*NORDUnet*, GÉANT	GÉANT
Storage				*DICE*, EUDAT
Compute	*ELIXIR*		*LUMI*	*DICE*, EGI-ACE
Data	*ELIXIR*, *ICOS*, EISCAT	*ESS ERIC*		COPERNICUS
Software				
Instrument		*ESS ERIC*		

Table 1

It can be noted that the existing examples of cross-border collaborations in the Nordic are in most cases addressing maximum one or two types of resources.

In the remainder of this chapter, a number of Nordic collaborations are presented, and their cross-border funding models discussed. When reviewing these, the above table and categorisation can serve as a guide to how the collaborations under discussion fit into the larger landscape, and to understand how different cross-border funding models can be applied. To aid the reader, the collaborations presented below are marked in *italics* in the table above.

It should be noted that the examples identified of federated e-Infrastructures are usually linked to Research Infrastructures on the ESFRI roadmap or an established Research Infrastructure ERIC. These are larger, cross-border research collaborations, involving many researchers in multiple countries, and often with significant budgets, with varying degrees of cross-border

funding. Federating resources for cross-border use appears to be a useful way to provide sustainable funding – cash or in-kind – for cross-border resources sharing in this case. Such domain-specific, federated resources have not been identified for the long tail of science. This reflects the learning from the EOSC Nordic demonstrators (chapter 4) that smaller resource collaborations depend on sharing national resources or sharing of resources funded by EC projects, both being short-term funding models.

3.2. Analysis of the funding models

A number of currently existing successful Nordic cross-border collaborations identified in 3.1 are analysed in detail in this chapter.

These collaboration cases have been selected because they address cross-border funding challenges in a successful fashion. Understanding the landscape structure/agent particularities of the funding mechanisms used and their selection criterion, is argued to provide valuable insight into guidelines for EOSC.

The selected cases can be clustered based on the business models adopted for acquiring resources and services. The following classification is aligned with the EOSC-hub deliverable *D12.3 Business models and procurement: evaluation and recommendations*[13]:

- ***Public procurement with demand aggregation and a Central Purchasing Body (NORDUnet)***: The demand for common services from research and educational users can be aggregated in order to get the best deal from suppliers in the market whilst complying with procurement regulations. The Central Purchasing Body (CPB) can be an EU body with a central role (e.g., GÉANT on behalf of the NRENs) or a corresponding national one (e.g., NORDUnet on behalf of its member or served institutions);
- ***Virtual Access via EC grant (DICE)***: Virtual Access (VA) refers to a specific financial instrument defined in the "European Research Infrastructures (including e-Infrastructures)" EC Work Programme, which is part of the Horizon 2020 framework program. The goal of this instrument is to reimburse the costs of service providers (also called "access providers") as beneficiaries of the H2020 grant for provisioning (via the internet) services to researchers. Virtual Access is similar to remote Transnational Access (TNA), but it does not allow differentiation between users. TNA requires a process to select users normally based on scientific excellence (e.g., for the access to a scarce resource such as a supercomputer).
- ***In-kind + in-cash cross-border pooling of resources (ICOS ERIC,ELIXIR, ESS, LUMI):*** The in-kind model is based on a community pooling their resources into a "hub" that encourages fairness of resource sharing. The in-kind contributions are also complemented by in-cash contributions, such as grant(s), EC, national or other, such as the ones supported by the ESFRI Roadmap process that can fund ESFRI RIs or ERICs in different phases, such as design phase, preparatory phase, implementation phase, etc. This combines the advantages of the in-kind model with those of dedicated funding

[13] https://documents.egi.eu/public/RetrieveFile?docid=3633&filename=EOSC-hub%20D12.3%20v1%20FINAL%20Public.pdf&version=2

such as EU/national grants, i.e., dedicated resources / personnel and concrete outputs, ultimately sticking to timelines for the construction/operation of the RI.

The analysis of the funding models takes into account four different main perspectives:

- The “procurer” perspective: meaning the perspective of the actors that are “procuring the resources” **(**please note that they are not necessarily the ultimate end users). This part is fundamental to understand who the different procurers are, how they procure and pay for the resources, and what triggers the demand;
- The “provider” perspective: it is important to clarify who owns the resources, how the provisioning of resources is organised and funded, and the cost drivers;
- The “strategic” perspective: this is usually the perspective of the funders or policy makers. This is essential to understand whether the mechanism is enabled by top-down regulations and policies or not.
- The end-user perspective: meaning those consuming the resources.

The main challenges and success factors have been analysed from all these different points of view.

The analysis has been structured according to the template described in Annex 1 and the cases are described in alphabetical order in the following sections. A summary of the main findings and lessons learnt of this analysis is documented in chapter 3.3.

3.2.1 Data Infrastructure Capacity for EOSC (DICE) project

Data Infrastructure Capacities for EOSC (DICE)[14]
Background information: The DICE project is a EU Horizon project like most others, but has a notable attribute of testing a specific funding model – Virtual Access (VA) – and is included here because of it. The DICE project brings together a network of computing and data centres, research infrastructures, and data repositories for the purpose of enabling a European storage and data management infrastructure for EOSC, providing generic services and building blocks to store, find, and access data in a consistent and persistent way. The majority of partners in the DICE consortium are also EUDAT CDI centers, representing all Nordic countries. DICE is funded by the European Commission under the Horizon 2020 project call INFRAEOSC-2018-2020. Through DICE, the EC enables storage providers around Europe to offer data storage facilities free at the point of use, to the research community through the VA funding mechanism. VA makes it possible for providers to be compensated by the EC while transparently offering new users their services at no cost – i.e. until June 2023[15].
Type of resources provided: DICE partners offer 14 data management and storage services associated with a total of more than 50 PB of storage capacity. Service categories include:

[14] dice-eosc.eu
[15] https://www.dice-eosc.eu/call-service-requests

- Personal/project workspaces
- Data archives
- Policy-based archives
- Data repositories
- Data discovery

Procurer perspective (meaning those that are **"procuring the resources"** and NOT the ultimate end-users)

Target procurer:
- The European Commission

How the resources can be procured
- The EC provides funding to partners in the DICE consortium (composed of major European computing and data centres). The funded organisations then ensure that the different services are available for request and use by end-users.

What triggers the demand
- Community managers or research infrastructures who mediate and recommend data management services for their end-users.
- End-users can have access to resources beyond their national borders

How the procurers "pay" for the resources
- The EC funding to each service provider under the Virtual Access funding mechanism. Each partner is expected to provide proof of consumption of the resources using an auditable accounting process.

Main challenges in procuring the resources
- N.A.

Success factors
- Service providers are willing to provide the services i.e. a structure and process exists via EUDAT
- The EUDAT services have been operating for a long time aimed for use across Europe.

Provider perspective

Who owns the resources
- DICE consortium partners (CINECA, CSC, FZJ, BSC, GRNET, SURF, KIT, MPCDF, IT4I, DKRZ, CESNET, GWDG, ETHZ, INFN, SNIC, SIGMA, DataCite, Cyl)

How the provision of resources is organised
- The majority of partners in the consortium belong to the EUDAT CDI. The resources are provisioned using the EUDAT workflow. For non-EUDAT partners, the resources are provisioned following each partner's existing processes.

How the provision of resources is funded

- Under the DICE project, the partners receive funding from the EC to provide the resources using the VA funding mechanism.

Cost drivers
What determines the costs of the resources:
- Labour costs for service operations and to set-up and customise services to satisfy end-users needs
- Infrastructure costs

Main challenges in providing resources
- Getting new users beyond the existing user base.
- The VA funding mechanism is not entirely clear
- Lack of clarity when aligning the existing accounting processes with the VA rules.
- Finite duration of the VA funding, as data services are typically needed for longer times than the DICE project duration (services mostly available on a pay-per-use basis after the VA funding ends).
- The process of providing and allocating resources is can be seen as unclear (especially the connection/integration with the EOSC Marketplace)
- The resources are being procured separately from other relevant resources such as computing.

Success factors
- An existing provision mechanism through the EUDAT network
- Available funding from the EC
- Sustained earlier collaboration within EUDAT CDI underpins the project

Strategic Perspective

Strategic drivers/motivations:
- The EC wants to promote open science and encourage cross-border provisioning of services.

Strategic or policy stakeholders:
- The European Commission
- EOSC related projects
- DICE consortium
- EUDAT

End User perspective (end users = those that consume the resources)

Description of end users:
- European researchers - under the DICE project, it's possible for e.g. an Italian researcher to use services from Germany.
- Generic services (as opposed to thematic services)

Main challenges:

- The end-users are mainly concerned about the continuity of service provision after the project ends (free during the project, but what happens after the project ends in terms of service fees?)

Success factors:

- The availability of free services for different research needs to pilot and experiment with new solutions
- The EUDAT services have been operating for a long time aimed at European use.

Other stakeholders perspectives

National funders: After the project ends, national funders can continue providing funding under the existing national schemes. That means that some users from other countries may no longer be eligible to use the resources after the project ends.
Research communities and research infrastructures: The services are often requested by communities or RI managers for their end-users (but they are also requested directly by the end-users). During DICE they have the opportunity to offer the DICE services for their end-users at no cost under the VA method. After the project ends, the communities or RIs may opt to enter into agreements with some of the service providers to continue using the service. In that case, they would become the actors "procuring the service".

3.2.2 ELIXIR

ELIXIR[16]

Background information:
ELIXIR is an intergovernmental organisation that brings together life science resources from across Europe. ELIXIR unites more than 220 of Europe's leading life science organisations in managing and safeguarding the increasing volume of data being generated by publicly funded research. It coordinates, integrates and sustains bioinformatics resources across its member states and enables users in academia and industry to access services vital for their research.

Type of resources provided:
ELIXIR resources include databases, software tools, training materials, cloud storage and supercomputers. The goal of ELIXIR is to coordinate these resources so that they form a single distributed infrastructure. This infrastructure makes it easier for scientists to find and share data, exchange expertise, and agree on best practices. ELIXIR's activities are divided into five areas called 'Platforms'. These are Data, Tools, Compute, Interoperability, and Training. The Platforms are managed by Platform leaders and the work is carried out by groups within the Platforms.
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 80 additional thematic databases and data deposition resources coordinated by ELIXIR and maintained

[16] elixir-europe.org

locally with ELIXIR partners. ELIXIR provides a registry of data resources through fairsharing.org run by the University of Oxford.
Tools: Analysing vast amounts of bioinformatics data requires specific software tools for high-throughput computation. Access to these tools (i.e. algorithms) are equally important as access to the data to both reuse the data and to validate research results. These tools are openly available through the ELIXIR partners.
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. That enables providing mechanisms for researchers to move their analysis to where the data is 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 Nodes22 and their partners, and by the European e-Infrastructures (i.e. EGI, EUDAT, GÉANT, etc).
Interoperability & standards: The Interoperability Platform aims at supporting the life science community in adopting standardised file formats, metadata, vocabularies and identifiers. That 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 made available. Thus ELIXIR provides tool kits, training materials and a technical platform for learning within the community.

Procurer perspective (meaning those that are **"procuring the resources"** and NOT the ultimate end-users)

Target procurers:
ELIXIR Communities bring together experts across Europe to develop standards, services and training within specific life science domains. The Communities also provide feedback on the Platform services, which helps ensure these services are practical and useful.

How the resources can be procured
Contributing membership fees to cover the operational costs of the respective national ELIXIR nodes.

What triggers the demand
The need for ([17])
- Efficient management of life science data
- Tools and training to handle the rising complexity of data
- A robust bioinformatics infrastructure
- Innovations and industry usage

How the procurers "pay" for the resources
Via membership fees

[17] https://elixir-europe.org/about-us/why-needed

Main challenges in procuring resources

Success factors

- Benefits of working together: (also known as relationship capital), we facilitate knowledge-sharing and cooperation.
- Bioinformatics resource uptake: we work to increase their usage and appreciation by users.
- Equal opportunity: we raise awareness of diversity and inclusiveness.
- Policy influence: we ensure that policy-makers are aware of the benefits of Open (FAIR) Science.
- Public awareness: we raise the public's awareness of bioinformatics and open science, including their socio-economic benefits.
- Research dissemination: our scientific legacy ensures increased awareness of developments related to research infrastructure and their resources.
- Research efficiency: we make infrastructure, resources and processes faster, easier to use, and more integrated.
- Research infrastructure sustainability: we work to increase its visibility and appreciation by funders.
- Skills development: (also known as human capital), we foster better skills for users and service providers.

Provider perspective

Who owns the resources

The ELIXIR Nodes. An ELIXIR Node is a collection of research institutes within a member country. Each Node has a lead institute that oversees the work of that Node.

How the provision of resources is organised

ELIXIR Nodes run the resources and services that are part of ELIXIR.

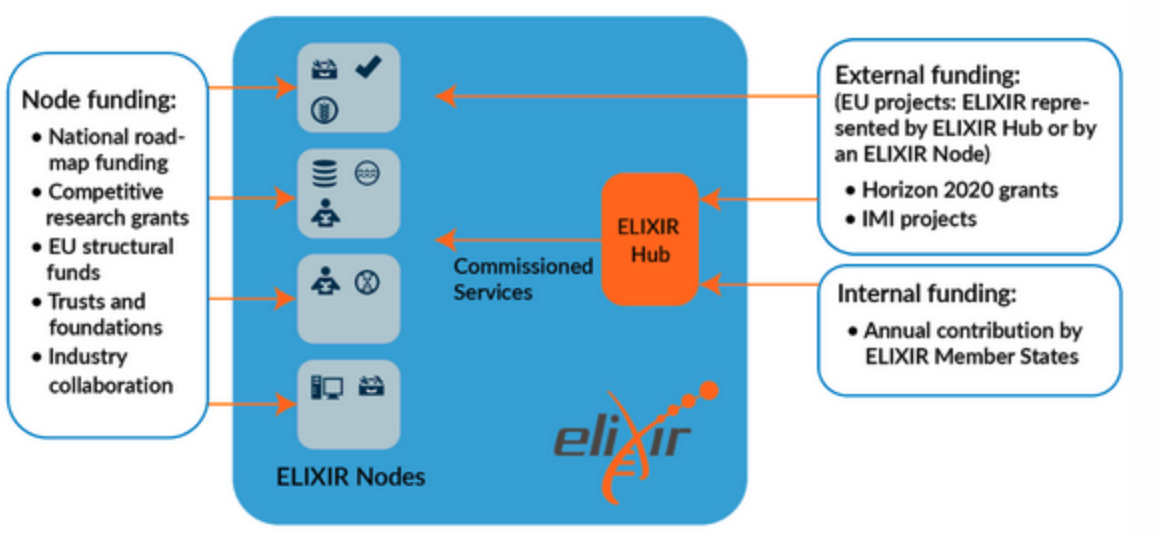
How the provision of resources is funded

As a distributed infrastructure, ELIXIR has a mixed funding model with contributions coming from several, primarily public, sources:

- The ELIXIR Hub is funded through membership fees paid by Member countries, and much of this funding is then transferred back to Nodes (e.g. via Implementation Studies) to deliver ELIXIR's five-year Scientific Programme. The participating countries contribute membership fees, proportional to the countries Net National Income (NNI), providing ELIXIR with a technical budget used to fund the development of shared services that support and link the nationally funded bioinformatics resources.
- ELIXIR Nodes are typically funded through national-level investments, supporting national coordination and the development and operation of services;
- ELIXIR Nodes also receive support from international funders, such as the US National Institute of Health (NIH);
- Finally, some ELIXIR Nodes can access European Union Structural Funds, for instance to support national coordination and the purchasing of facilities and grant

funding from the European Union, under Horizon Europe and the Innovative Medicine Initiative (IMI).

Other sources of funding for Nodes include foundations (e.g. the Wellcome Trust and Wallenberg Foundation), and also the industry, though the latter is modest compared to the public funding received by Nodes.



Cost drivers
See types of resources provided

Main challenges in providing resources
Research infrastructure sustainability: ELIXIR work to increase its visibility and appreciation by funders.

Success factors
The Nodes guide the direction of ELIXIR in alignment with their national priorities
ELIXIR connects national resources into transnational infrastructure. Activities are organised into a five-year programme in consultation with the Heads of Nodes and the ELIXIR scientific advisory board

Strategic Perspective

Strategic drivers/motivations:
The EC and the Member States recognise the importance of handling and analysing the massive amounts of data now generated in life science to address grand societal challenges (e.g. the COVID emergency). ELIXIR's five-year programmes, 2014-2018 (completed), 2019-2023 (ongoing), 2024-2028 (under development) are introduced by the ELIXIR director and approved by the ELIXIR board consisting of ELIXIR member state representatives.

Strategic or policy stakeholders:
- Member States
- European Commission

• International funders
End-user perspective (end users = those that consume the resources)
Description of end-users: Global researchers ELIXIR is committed to Open Access as a core principle for publicly funded research. ELIXIR Core Data Resources should reflect this commitment and have terms of use or a licence that enables the reuse and remixing of data. The Creative Commons licenses CC0, CC-BY or CC-BY-SA are all conformant with the Open Definition (http://opendefinition.org/licenses/), as are equivalent open terms of use.
Main challenges: Research infrastructure sustainability
Success factors: Open Access as the driving principle

3.2.3 LUMI

EuroHPC[18] / LUMI[19]
Background information: The European High-Performance Computing Joint Undertaking (EuroHPC JU) is a European partnership in High Performance Computing (HPC), enabling the pooling of European HPC resources - financial and human - to deliver a pan-European HPC infrastructure. EuroHPC has a budget of €1 billion, half from the EU and half from Member States (MSs), but varying much from state to state depending on willingness and ability to engage in the overall initiative. Roughly half of the total EuroHPC budget is financing three very large and several smaller HPC installations, situated in specific host MSs. Typically the host MS finances most if not all the non-EU part, through a consortium consisting of at least one MS as well as the EU. One such consortium is LUMI which includes 10 countries jointly financing HPC hardware, operations and support to deliver one of the world's most powerful supercomputers, hosted in Finland. What makes the LUMI consortium special and different from the rest of EuroHPC is the amount of co-owner MSs as well as the degree of budget distribution across MS partners. Many MSs have made a significant cross border investment, which is not the case for EuroHPC otherwise, or at least nowhere near the same degree. The non-LUMI part of EuroHPC is less interesting for our purpose here since the resources are on the one hand jointly owned, but also strictly distributed back to investors (owners), meaning that only the EU-owned 50% part is a truly joint European infrastructure.
Type of resources provided: • HPC - i.e. CPU clock cycles, including system administration

[18] eurohpc-ju.europa.eu
[19] lumi-supercomputer.eu

- High-level user support (modelling, simulation, data analytics software development, tuning and installation)

Procurer perspective (meaning those that are **"procuring the resources"** and NOT the ultimate end-users)

Target procurers: LUMI consortium members & EC:
- National research ministries
- Universities
- Computer Centres

it depends on the different countries

How the resources can be procured
The LUMI HPC facility was initially procured by states only (national computer centres) and only with agreement within the owner consortium as well as the EuroHPC Governing Board.
Actual hardware procurement is through tender run by the host in cooperation with and mandated by the EuroHPC JU Governing board.
Researchers and research groups *cannot* procure but only apply for grants through competitive calls or strategic allocation. There are national calls (for the national shares) and EU calls (for the EU share). Each country has its access/allocation policies regarding its own share.

What triggers the demand
The ever-growing need for computational resources is driven by the increase of computing needs (for modelling, simulation and data analytics) as an important if not fundamental scientific method.
HPC budgets that the MSs are willing to invest are therefore increasing, and the rationale of value for money (lower TCO), the economy of scale, and pooling a critical mass of competence is gaining momentum.

How the procurers "pay" for the resources
EU member states (or national computer centres) pay cash.

Main challenges in procuring resources
Negotiating the financial contribution of the different countries and the associated resources.

Success factors
Willingness for participating MSs to invest in critical strategic infrastructure in only one country.

Provider perspective

Who owns the resources
LUMI is financed with €202 million through a consortium composed of Finland, Belgium, Czech Republic, Denmark, Estonia, Iceland, Norway, Poland, Sweden, and Switzerland. The EU has full right of disposal to 50% and distributes that resources through competitive call, like seen from PRACE.

How the provision of resources is organised
Joint provision of resources (public call for tender) managed by the owner consortium and overseen by the EuroHPC Governing Board.

How the provision of resources is funded
Ownership of the resources is distributed one-to-one according to financial contribution. The EU owns 50% of all, and the Member States own according to the percentage of Total Cost of Ownership (TCO) they have financed.

Cost drivers
The costs of the resources (TCO) is determined by a number of complex yet very important factors:

- cost of HPC hardware, where economy of scale is significant
- cost of electricity for HPC operations and cooling, where electricity is the major operational cost amounting to 20 to 40% TCO, and the cost of electricity is significantly different across European geographical price zones.
- Labour cost, with significant difference across European borders, yet only a small fraction of TCO.
- Operations and system administration (small)

Main challenges in providing resources
Skepticism from the providers to engage in a provisioning model very different from the traditional one (local vs cross-border provisioning). Providing resources cross-border can undermine the local interests, reputation and business of national service providers.

Success factors
The paramount success factor is the ability to successfully deliver much more HPC at much lower TCO.

Strategic Perspective

Strategic drivers/motivations
The strategic perspective is that EuroHPC in general but LUMI in particular can, and indeed has, changed the European HPC landscape. The duplication of effort and extreme loss of potential economy of scale and gathering of critical competence mass, has been changed and partly overcome.

The time when every university needed to have an HPC installation, and every country needed to have flagship HPC is partly over. The need to deal with growing HPC budgets facilitates an increased need to focus on value for money (lower TCO), economy of scale, and pooling critical mass of competence.

Strategic or policy stakeholders:
National research ministries and some (few) larger computer centers see the possibility for increased capability (more resources for the same budget) through international joint investments - as is also the case with ESFRI.
There are still some cases where campus/national HPC organisations perceive this model as a threat for their business because they feel that are moving from having full local control

over (qualitative/quantitative inferior) resources to having lesser or no control over superior resources.
End User perspective (end users = those that consume the resources)
End user description: Research groups, i.e. national and international grant holders, as well as larger well organised strategic communities (ESFRI-like)
Success factors: The national and European researchers get more and better HPC (CPU clock cycles) as well as higher levels of support.

3.2.4 European Spallation Source (ESS) ERIC

ESS - European Spallation Source[20] ERIC
Background information: The European Spallation Source ERIC is a research infrastructure under construction in Lund, Sweden, that will become the world's most powerful pulsed neutron source. The Data Management and Software Centre (DMSC) is located in Copenhagen, Denmark.
Type of resources provided: ● In ESS, Sweden and Denmark share competence so that the Neutron Source, the experiment stations, and all technical support and logistics are currently built in Lund, Sweden. The ESS is built as a single facility located in two countries, operated by a single organisation (rather than having separate, smaller facilities and multiple organisations). ● The ESS Data Management and Software Centre (DMSC) is located in central Copenhagen, Denmark, while the core instrument facilities are in Lund, Sweden.
Procurer perspective (meaning those that are **"procuring the resources"** and NOT the ultimate end-users)
Target procurers: ● Research-performing organisations
How the resources can be procured The installations are under construction.
What triggers the demand At the very beginning, ESS was a bottom-up project, driven by scientific interests in building the next generation neutron source - from the OECD to the European Task Force to ESS Scandinavia.
How the procurers "pay" for the resources In ESS the construction costs, budgeted at EUR 1,843 billion, Sweden and Denmark are responsible for approximately half. The remaining construction costs are covered by other

[20] europeanspallationsource.se

partner countries. In this case the cost sharing is agreed as part of the founding of the collaboration.
Main challenges in procuring resources The Nordic countries have traditionally not opted to drive the establishment of large scale RI within the region. Much more commonly the Nordics have opted for these to be built in the larger countries and joining as non-hosting members. In part because scientific communities in absolute numbers are small and because the financial responsibilities are very large compared to the overall science budgets. The Nordics also tend - or for that reason - to have quite a stringent scientific focus when making a decision to fund an RI.
Success factors ESS is an example of how regional cooperation between a couple of Nordic countries can initiate a bigger and wider collaboration covering relevant parts of the whole of Europe.
Provider perspective
Who owns the resources The ESS Data Management and Software Centre (DMSC) is located in central Copenhagen, Denmark, while the core instrument facilities are in Lund, Sweden.
How the provision of resources is organised The resource is under construction (Science & Instruments | ESS (europeanspallationsource.se))
How the provision of resources is funded In ESS the construction costs, budgeted at EUR 1,843 billion, Sweden and Denmark are responsible for approximately half. The remaining construction costs are covered by partner countries. In this case the cost sharing is agreed as part of the founding of the collaboration.
Cost drivers Innovative science - Most existing neutron sources are based on nuclear reactors, an approach that has reached its maximum capability and cannot be developed further.
Main challenges in providing resources The resource is under construction (Science & Instruments | ESS (europeanspallationsource.se)) As the project gained political momentum, concerns were being raised about it within the broader scientific community - specifically on the size of the user base and the scientific value of the resource - and, perhaps most important, concerns about the cost-benefit ratio of ESS. Behind those concerns were two issues: 1. Would the establishment of ESS reduce resources for other scientific fields in a situation of fixed national budgets? 2. Who bears the financial risks for the construction costs or increased operational costs in the long run?

These concerns were born out of the past experience regarding competition for funding between scientific fields and Sweden giving ESS priority over MAX IV laboratory[21].
Strategic Perspective
Strategic drivers/motivations: At a political level, the project reached momentum by being framed as not only a special tool for a small group of scientists but a tool that would be interesting in a wide scientific application and also with industry, and from an industrial policy perspective to support the narrative of a regional cluster of science and industry, supported by economic arguments about the importance of ESS in terms of influx of highly skilled people to the area.
Strategic or policy stakeholders: It is reasonable to suggest that the drive to put ESS in Lund was motivated more by regional industrial and political interests rather than scientific needs. There was certainly a great deal of criticism of the decision. Already in 2008 the Swedish Royal Academy of Science addressed the Swedish government in a letter, criticising the preconditions for ESS, namely, the scientific case, the user base and the economic sustainability of ESS, suggesting that the establishment of MAX IV would be more beneficial for Swedish science. Similar criticism was raised in Denmark.
End User perspective (end users = those that consume the resources)
Description of end users: neutron scientists
Success factors: 1. Science has moved into the political arena, as a tool to support other policy goals - industrial and regional development goals. 2. The formalisation of a decision making structure through the establishment of ESFRI and its roadmap. Not only did ESFRI create a decision structure that enabled the ESS Scandinavia bid, it worked on the same logic of science supporting industrial policies and geopolitics. But the financial risk of such an undertaking is still significant for small countries. In comparison the current Danish share for the construction of ESS is larger than the state budget for public research grants. Consequently also the political backlash of budget overruns. ESS Scandinavia was established to share the financial risks.

3.2.5 Integrated Carbon Observation System (ICOS) ERIC

ICOS ERIC[22]
Background information: ICOS is a research infrastructure providing quality data and derived products. Data is provided under CCBY 4.0 license.

[21] maxiv.lu.se
[22] www.icos-cp.eu/about/organisation-governance/icos-eric

The RI is coordinated by ICOS ERIC which includes Head Offices and Carbon Portal. In addition to ICOS ERIC the RI consists of three thematic centres, Central Facilities, providing services like instrument testing, network support, training, development of methods and sensors, analytical services etc. and National Networks which provide data to the RI.
Type of resources provided: ● Free access to high-quality and standardised greenhouse gas data, as well as to scientific and educational products and services ● Instrument testing ● Network support and training ● Development of methods and sensors ● Analytical services
Procurer perspective (meaning those that are **"procuring the resources"** and NOT the ultimate end-users)
Target procurers: ● National research ministries as members or observers of ICOS ERIC ● Universities and research institutions as part of the Central Facility consortia or as part of the National Network providing the data
How the resources can be procured Everything is financed by countries that are members or observers of ICOS ERIC or third parties like JRC which join ICOS ERIC with a separate agreement.
What triggers the demand Scientific needs for quality data and derived products.
How the procurers "pay" for the resources No payment data is provided free of any charge under CCBY 4.0 license.
Success factors Data is provided efficiently and used.
Provider perspective
Who owns the resources Resources are owned by those creating them so basically research performing organisations in different countries and Central Facilities and the National Networks give ICOS ERIC very wide user rights.
How the provision of resources is organised ICOS ERIC is governed by its statutes and the Central Facilities and the National Networks are linked to the RI through agreements with ICOS ERIC.
How the provision of resources is funded The RI is financed by countries that are members or observers of ICOS ERIC or third parties like JRC which join ICOS ERIC with separate agreement.

Financing is done according to the following principles:

- Resourcing of the ICOS National Networks is organised nationally.
- ICOS Central Facilities, operated outside the ICOS ERIC either as national or as multinational consortia, are funded in major part from the hosting countries and in lesser part by ICOS ERIC through reallocation of annual contributions.
- ICOS ERIC integrated activities are funded by annual contributions and by host premium contributions.

Annual membership contribution to ICOS ERIC is based on the following variables:

- common basic contribution (50 % of the common contributions),
- common GNI-based contribution (50 % of the common contributions),
- station-based contributions.

Hosting countries are committed to pay host premium contributions to ICOS ERIC (Head Office, Carbon Portal).

As a landmark RI under the ESFRI, the RI also has the possibility to apply for funding money from dedicated infra projects under EU funding to develop operations.

National funding comes from national programmes.

Cost drivers

- cost of developing and maintaining services and tools
- labour cost

Main challenges in providing resources
Changes in national funding and national roadmaps.

Success factors
Quality data and derived products are provided efficiently with good, well organised access.

Strategic Perspective

Strategic drivers/motivations: Open science and open access to quality data

Strategic or policy stakeholders: EU and countries

End User perspective (end users = those that consume the resources)

Description of end users: Anyone interested, mostly scientists

Main challenges: Address why to use ICOS labeled data from Carbon Portal instead of getting data straight from the research performing organisations.

Success factors: Amount of use and science made based on the data.

3.2.6 NORDUnet

NORDUnet[23]
Background information: NORDUnet is the joint, international network of the Nordic national research and education networks (NRENs). NORDUnet was started as a technical collaboration between network engineers of the Nordic academic sector and developed into a provider of international connectivity for the Nordic NRENs in the second half of the 1980's. NORDUnet came to represent the Nordic research and education sector in European and global network collaborations and evolved into a platform for network-centric services for the academic sector. NORDUnet has existed as a Nordic research infrastructure collaboration for 40 years, and has been successful in delivery of infrastructure, in increasing Nordic influence on the evolution of European and global research infrastructures, and has been largely stable in terms of cost, cost sharing, governance, and service delivery despite dramatic changes in the infrastructure landscape.
Type of resources provided: Network resources and demanding, network-centric services where scale is an advantage (i.e., Zoom for Nordic R&E)
Procurer perspective (meaning those that are **"procuring the resources"** and NOT the ultimate end-users)
Target procurers: National Research and Education networks in the Nordic countries (>80% of the turnover) plus select services for cross-border research collaboration (i.e., Nordic WLCG Tier-1) and international R&E networks (i.e., GÉANT, European NRENs).
How the resources can be procured The core services of NORDUnet are *cost shared*. The cost of the network and other core services are divided among the owners (the five Nordic NRENs, who are also the customers) according to the cost distribution key of the Nordic Council of Ministers (largely proportional to GPD). Some network-centric services are provided on a cost-recovery basis, based on resource consumption. This mechanism is used for services that may not be equally consumed across the owners. In this case, cost recovery is agreed by the NORDUnet board of directors (representing the owners). Cost recovery is always with a national R&E network, not with the eventual end user.
What triggers the demand NORDUnet is the single provider of international network capacity for the Nordic R&E networks and by extension Nordic R&E institutions. Demand follows the network traffic usage of the researcher education practices. Network-centric services are provided where the owners agree that there is strategic or scale advantage in offering such services in a centralised manner rather than nationally, at the

[23] www.nordu.net

institution level, or procured from the market. Such decisions are based on national R&E ICT strategy, service demand of the academic sector, and open market conditions. Once a service offering is agreed, demand is driven by service consumption at the national level, through the national provider (NREN).
How the procurers "pay" for the resources Cost shared services (the majority of the cost) is agreed as part of the annual budget and cost shared. Cost recovery services are invoiced quarterly. Since NORDUnet sells services only to national providers, who in turn re-sell to end-users, the number of financial transactions are limited. The majority of cost-shared services (the core network) are offered in a best-effort manner through over-provisioning. There is no individual accounting or payment for resource consumption.
Main challenges in procuring resources For the Nordic NRENs, the main challenges are strategy and cost. For the core infrastructure, cost issues are essential and are addressed through technology choices and federation of national resources. Cost is driven by strategic choices about level of service and global reach of infrastructure. For network-centric services, strategic choices centre on what is best done at institution, national, regional level, or acquired from the market. This challenge is addressed through NORDUnet governance and in collaboration among the stakeholders.
Success factors NORDUnet enables the Nordic countries to jointly own and control an advanced, global network infrastructure and to provide key services such as Zoom at scale. NORDUnet also enables the Nordic countries to have a European and global role - as a driver of Research Infrastructures strategy and in European governance fora - larger than what smaller countries would otherwise have.
Provider perspective
Who owns the resources The majority of the resources offered through NORDUnet are owned by NORDUnet, acquired from the market through public procurement. In some cases, resources are acquired from European e-Infrastructures (i.e., GÉANT) or are federated from the owners (i.e., network resources of the national Nordic networks).
How the provision of resources is organised NORDUnet is a limited company, governed by a board of directors representing the five Nordic R&E networks, who are also the major customers and consumers of resources. The board is elected by the owners, the Nordic ministries or research and education.
How the provision of resources is funded NORDUnet is funded by the Nordic research and education network. The majority of the funding comes through a cost-sharing mechanism based on GPR of the member countries; remaining funding comes through cost-recovery for services provided to NRENs and a few, major Research Infrastructures. NORDUnet does not receive national or EC funding for resource provision.

Cost drivers NORDUnet cost is primarily the cost of core infrastructure - cost of network capacity, network equipment, servers and storage for network-centric services.
Main challenges in providing resources Key challenges are stable, long-term strategy and capacity planning for core infrastructure, and clarity of strategy and short-term evolution of network-centric services.
Success factors NORDUnet is stable - governance, organisational, financial - and largely has a clear mission and clear mandate. NORDUnet is recognised and trusted in the Nordic countries, in Europe and globally. Effort is made to ensure joint Nordic positions on strategic issues, enabling NORDUnet to act and represent with clarity.
Strategic Perspective
Strategic drivers/motivations: For 40 years, the main strategic drivers have been digital sovereignty and cost efficiency. It is key to the Nordic NRENs to be an equal player on the global scene and be able to independently make decisions on issues impacting future evolution of research infrastructure.
Strategic or policy stakeholders: Nordic research and education networks (NRENS), who are also the main governance. Nordic ministries of research and education, who are the owners. Other Nordic Research Infrastructures. Major Research Infrastructures based in the Nordic countries.
End User perspective (end users = those that consume the resources)
Resources are exploited by both research projects, Research Infrastructures, and individual researchers, educators, and students in the academic sector.
Success factors: The provisioning, cost, payment, and governance is largely invisible to the end users.

3.3. Essential lessons learnt

From the analysis of the six cross-border collaborations in chapter 3.2 the following findings and lessons can be drawn:

A clear difference can be noted between cross-border collaboration funding models for generic and discipline-specific e-infrastructures

The EuroHPC/LUMI (i.e. generic e-infrastructure) funding model differs from the funding models of discipline-specific infrastructures (e.g. ESFRIs and ERICs) for two main reasons:

- EuroHPC/LUMI is strongly top-down driven, for negotiation as well as finance, whereas ESFRIs and ERICs tend to grow from the bottom through international scientific community consolidation, financed primarily by research funders or their ministries;
- EuroHPC/LUMI provides a generic infrastructure used by and relevant for many scientific disciplines. Hence, in contrast to ESFRI, financing comes not from research funder budgets but dedicated e-Infrastructure provider organisations or their ministries.

EuroHPC/LUMI might resemble an ESFRI or an ERIC only from a ministry point of view.

The financial models differ depending on the type of resources provided and the type of funding mechanism

NORDunet has a shared infrastructure and they have a model to distribute the costs - the cost-sharing model is neither based on the actual costs nor on actual usage. Each member pays their share of total cost of ownership according to national GDP as a percentage of total GDP of member countries. I.e. the bill also changes every year according to the GDP. The main benefit is that it allows the countries to procure a more advanced network, with a slight overcapacity by aggregating the demand, without needing to calculate, control and distribute costs on usage. Only total usage is seen as relevant.

Virtual Access (VA) is an example of opening up existing resources to more researchers via public funding. The goal of this financial instrument is to reimburse service providers the cost of provisioning services to researchers via the EC grant (service providers are grant beneficiaries and can claim costs, either via unit costs, actual costs, or a mix of both). This is a good method for

1. service providers to be reimbursed;
2. end-users to access services that would otherwise be out of reach;
3. the EC to fund and support open science practices across Europe.

However, VA is quite complex and there are still unresolved accounting issues. VA can work well with on demand resources (e.g. computing, storage resources). Applying VA to data services seems to be more complicated.

Strategic and political commitment, especially at national level, is fundamental for driving successful cross-border collaborations

One of the most significant lessons learnt from the LUMI case is that even with a very clear and convincing rationale, the cross-border collaboration was only possible due to high-level strategic understanding and commitment, including top-down government engagement. Initially, many HPC centres, and national actors with vested interests, may have felt threatened by the LUMI rationale, which could be seen as some competing facility. The political driver has been very important for bringing a cultural change.

In the case of ERICs and ESFRIs, the main drivers are the national and ESFRI roadmaps addressing the importance of a certain RI for a specific scientific field. The funding model relies

mainly on country contributions or membership fees from the participating research-performing organisations. Country contributions are made possible through national research groups' ability to obtain significant national research funding. The downside of the ERIC model is that ERICs are very much dependent on a country's commitments to maintain the services in the long term. If the priorities of the country are changing, the funding might too.

Also in the case of ESS, the political commitment as a driving force to form a major cross-border collaboration was a key factor bringing together networks of institutions in Scandinavia.

Cost-sharing of best-effort infrastructure works if there is a clear scaling advantage

LUMI has been motivated by the need to deal with growing HPC budgets, prompting an increased need to focus on value for money (lower TCO), the economy of scale, and pooling critical mass of competence. EuroHPC in general but LUMI, in particular, can, and indeed has, changed the European HPC landscape. The hitherto duplication of infrastructure and support effort has been restructured. A significant economy of scale and gathering of critical competence mass has been accomplished by overcoming the otherwise lacking ability to fund cross-border HPC infrastructure and support resources.

The economy of scale is also one of the main drivers of NORDUnet, enabling the Nordic countries to jointly own and control advanced, global network infrastructure and to provide key services such as Zoom at scale. Network-centric services are provided where the owners agree that it is a strategic or scale advantage in offering such services in a centralised manner rather than nationally, at the institution level, or procured from the market. Such decisions are based on national R&E ICT strategy, service demand of the academic sector, and open market conditions. Once a service offer is agreed upon, demand is driven by service consumption at the national level, through the national provider (NREN).

The added value of the cross-border collaboration must be clear. It is the only way to engage stakeholders.

ELIXIR, ICOS, LUMI and NORDUnet, all allow the stakeholders to get something they could not get alone. That this must be the driver of a collaboration seems obvious, but it is not always clear. Hence, it is fundamental to have a clear value proposition brought by cross-border collaboration. However, it often happens that value propositions do not state who gets said value, and who does not. I.e. national value might not be value to a specific person, group, or institution. To the contrary, it might be a threat due to a shift in strategic focus or funding. Therefore, value propositions need to be backed not only by economic commitment, but also political will and power.

While in the case of LUMI, NORDUnet and ELIXIR the value relies on the capability of scaling up resources, in the case of ICOS ERIC and the ESFRIs the added value is seen in the countries to facilitate science in a specific field.

In all cases, the understanding of the added value must be from an owner point of view, including ability and willingness to also finance. In sum, cross-border collaboration must have

national anchoring in terms of defining added value as well as contributing to financing. In current European cross-border e-Infrastructure collaboration, this is surprisingly seldom the case.

Cross-border collaborations on a smaller scale (e.g. at regional level) empower the individual members of the collaboration and reduce their financial risks

NORDUnet for example enables the Nordic countries to have a European and global role - as a driver of Research Infrastructures strategy and in European governance fora - larger than what smaller countries would otherwise have.

ESS is a good example of how regional cooperation between some Nordic countries can initiate a much bigger and wider collaboration covering relevant parts of the whole of Europe. In the case of ESS, the established cross-border collaboration was also a way to reduce the financial risks of the individual countries that for such a massive undertaking is still significant for small countries. In comparison, the current Danish share for the construction of ESS is larger than the state budget for public research grants. Consequently also the political backlash of budget overruns.

LUMI member countries are individually all politically not very strong, but collectively instantly become a very strong political force in the European landscape. LUMI is also a clear case of political empowerment and lowering of financial risks, which is especially clear for the smaller northern European countries at national level.

Virtual Access could be a good mechanism to widen the user base across borders

The main benefit of the Virtual Access funding instrument is that it provides flexibility and freedom for service providers to serve users that would otherwise be out of reach due to national funding constraints. Hence, the Virtual Access funding mechanism could be a mechanism to support cross-border service provision across Europe.

In the DICE project, for example, the Virtual Access instrument allows the consortium partners to provide data management and storage services to basically any European researcher. With Virtual Access being a fairly new funding instrument, some of its rules and requirements are unclear and may be difficult to interpret.

4 Lessons learnt from the EOSC Nordic demonstrators

The subject of *funding models* specifically for research e-Infrastructure is in many ways different from mainstream research funding. This difference affects most research and research groups in general and those dependent on or aiming for expensive and complicated e-Infrastructures in particular. The EOSC Nordic demonstrators[24] comprise research groups with different degrees

[24] https://www.eosc-nordic.eu/demonstrating-eosc-nordic/

of e-Infrastructure needs that are mostly suited for being provided within international cooperation.

The EOSC Nordic WP2 effort has looked into the needs of these research groups in terms of cross-border infrastructures focusing on the existing challenges for their adoption.

Challenges and blocking factors are investigated, followed by discerning the potential paths forward for these demonstrators.

1. **Clarifying the challenges**
    - Research groups working across borders need mechanisms to secure joint infrastructure for computing and storage. These mechanisms must extend beyond short-term or ad-hoc project funding. There is a need for long-term resources that can be accessed and used in a cross-border setting, by cross-border research collaborations and consortia, for example for long-term research data catalogues. The archeology demonstrator is an example with such needs. Furthermore, there is a requirement for such mechanisms to be accessible and usable by research groups of a scale below that of ESFRI projects and major, strategic research collaborations.
    - There is a need for funding or cost-recovery models for the delivery of resources to research collaborations as mentioned above, such that institutional, national or European service providers can deliver to these needs and recover cost from collaboration partners or funding bodies. Both cross-border cost recovery and models for jointly operated infrastructures (such as LUMI) may offer solutions to this challenge.
    - Currently, most funding models for e-Infrastructure are not suitable for cross-border funding or cost recovery due to the following limitations:
        - University libraries and IT departments mainly fund basic and generic infrastructures. Faculties and others typically do not fund e-Infrastructures at all. As a result, institutional infrastructures cannot be counted on to provide resources for cross-border research. The Archeology use case is an example of such resources not being available
        - National e-Infrastructure providers mostly offer resources only for national use. There is a lack of models for cost-recovery from cross-border users and incentives for offering such services. Creating joint, cross-border resources to meet the requirements is challenging in the face of different national funding cycles, strategic priorities, and governance and financing models. So far, few such infrastructures have been created. Overcoming this challenge, by funding cross-border activities or letting European researchers access valuable resources funded nationally, will offer a significant contribution to scientific excellence, new markets, recognition and visibility beyond national borders..
        - National grant funding agencies, public and private, may most commonly fund research and domain-specific instruments and facilities. Such funding is seldom for community-wide e-Infrastructure, being much too expensive and complex for single-group project funding.
        - National roadmap initiatives do not fund infrastructure in all countries.
        - European Strategy Forum on Research Infrastructures (ESFRIs) and the European Research Infrastructure Consortiums (ERICs) enable cross-border funding of instruments and domain-specific e-Infrastructure, mainly using national resources with little EU funding. However, these mechanisms are by their nature

limited to structured research communities. To them, subject-specific e-Infrastructures are in focus, not generic e-Infrastructure.
    - Current national funding models lack European coordination, and may lead to duplication of local and national e-Infrastructures. These funds are also often not sufficient to meet internationalised requirements, affecting the quality of the implemented solutions.
- Lessons learnt and a possible way forward:
    - It is important to enable the use of national funding for community-wide, generic, e-Infrastructure, across EU borders, and to combine with EU funding (via procurements or calls). This could be done in one or more of the following manners:
        - Ensure that infrastructure for long-tail research collaborations is provided and that funding or cost-recovery models allow cross-border use. This may require changes to mandates and priorities for the stakeholders involved.
        - Facilitate the use of resources from national e-Infrastructure provider organisations across borders, and enable national and EC funding to be used for cross-border resource funding.
        - Ensure adequate funding for generic community-wide e-Infrastructures. Do not expect national grant funding (research councils) to finance such e-Infrastructures. It is not their mission.
        - Increase funding of national roadmap initiatives, and encourage them to support cross-border initiatives.
        - Support and encourage ESFRI's and ERIC's, as those models appear to be working well for domain specific, cross-border e-Infrastructure.
        - Dedicate EU funding to matching national financing. I.e. abolish subsidising of national e-Infrastructure provider organisations.

5 Recommendations

Based on the lessons learnt and findings coming from the analysis of the funding models used in the cross-border collaborations at Nordic level and the analysis of the challenges encountered by the research groups working on the EOSC Nordic demonstrators, this chapter elaborates a set of recommendations that could contribute to the discussions on the definition of a funding model suitable for EOSC.

The recommendations below reflect the views of the EOSC Nordic consortium partners.

EOSC aims to enable the open sharing of knowledge and the re-use of research outputs[25]. To do so it is aiming at federating generic and subject-specific infrastructures pulling together a broad set of resources and services that will be offered to European researchers.

Recommendation 1: **Trying to identify a single funding model for EOSC is very challenging and not an appropriate solution. Breaking EOSC down into different** (more

[25] https://data.consilium.europa.eu/doc/document/ST-14308-2021-INIT/en/pdf

clearly defined) **components that can be aggregated and provided internationally, according to different strategic rationales and different funding schemes would be the best strategy to move forward.**

Our analysis shows that there exists several successful but quite different funding models, applied to different resources (data, network, storage, compute, etc.) and different purposes (generic versus subject-specific e-Infrastructures). A model that works well for subject specific data (typically the ESFRIs and the ERICs) does not necessarily work for other resources such as generic networks or computing (typically the e-Infrastructures provider organisations).

Therefore, a better acknowledgement is recommended, that different types of e-Infrastructures live in different funding realities (e.g. subject specific vs. generic) and need different types of funding models – as is outlined in section 3.1 “Types of collaborations” with the 2-dimensional table.

Recommendation 2: **Identifying more clearly what is in the scope of EOSC (type of resources and value proposition of offering) can facilitate the identification of appropriate funding models**

The definition of the resources and services that will be part of the EOSC Core and EOSC Exchange is still very vague. Without a clear definition of what EOSC offers and to whom, it is very difficult to identify suitable funding models.

Types of funding models, based on the 2-dimensional table in section 3.1 “Types of collaborations” is a starting point, but for EOSC offerings such as EOSC Core and EOSC Exchange need to be further clarified, and also broken down into manageable infrastructure components. Owners (interested national parties and/or researchers that are willing to co-fund) of said needed infrastructure components must be identified. Only then can a discussion occur, of who can best provide, based on which funding model. In all the established cross-border collaborations in the Nordics, this has been the case: national funding has been found, and pooled internationally.

The current catch-all funding model for EOSC infrastructure, i.e. EU calls within Horizon Europe, does not suffice.

Recommendation 3: **EC funding for EOSC should be substantially complemented by national funding linked to the specific investments, preferably articulated in connection to national strategic interests.**

A cultural change in Europe is necessary to move from the established concept of national funding to national provider organisations, supporting the provision of national generic resources, towards the concept of national funding supporting cross-border provisioning of joint

European resources advocated by EOSC, or other pan-european collaborations, in order to facilitate economy of scale, and to challenge entrenched interests and actors.

To better understand the current lacking or limited cross-border national funding for e-Infrastructure, it is recommended that the nature of the special e-Infrastructure provider organisation is better understood, as a political and economic entity, with specific local, national and international restrictions, abilities, competences and motivations – including an inclination to be conservative towards the entrenched interests.

Willingness to invest in cross-border initiatives is currently not strategically appreciated, thus not enforced top-down. LUMI and NORDUnet are exceptions in the current landscape. This trend leads to the local duplication of e-Infrastructure solutions that could otherwise be purchased and provisioned cross-border.

ESFRI/ERICs have also overcome the traditional concept of local/national provisioning of resources because their infrastructural needs are mainly driven by the research agenda of specific communities. The researchers are an integral part of the system driving these national investments cross-border. This has been possible because of the subject-specific nature of ESFRIs. This could be a challenge for EOSC given its generic nature.

For EOSC to succeed, its funding models must be articulated in connection with national strategic interests and national cash funding. Further, for cross-border collaborations, such as EOSC, it is essential that funding is functionally allocated to relevant projects and actors.

This will also reduce investments in isolated projects with little national strategic anchoring and no sustainable funding contributions.

Recommendation 4: **A working cross-border funding model for generic infrastructures must be found. The LUMI and NORDUnet funding models can serve as starting points.**

Generic infrastructures, such as those offered by national e-Infrastructure providers, are essential. Many researchers do not fit into the established ESFRI roadmap landscape, either because their research is too specialised to be on the roadmap or does not have sufficient national focus to be prioritised nationally. Likewise, some research collaborations are shorter-lived and will not fit the timeline of the ESFRI roadmap. This long tail of science is a significant part of the research landscape, and must be served. Generic infrastructures meet that requirement, and are often provided at university and/or national provider level, but not at the cross-border level.

LUMI and NORDUnet offer two different funding models for cross-border generic resources. Both models work for their specific domains. Both models work because of scaling advantages (economic, political as well as research impact). This topic should be further explored in other funding models for generic resources, for the strong research communities as well as the long tail of science. In NORDUnet, the model works by limiting the funding to a generic, best-effort resource. In LUMI, the model works by ensuring that each partner retains control over a fraction of the resource corresponding to their investment. The long tail of science, however, cannot rely

on these examples, as they will have a hard time getting the top-down political and economic engagement seen in the NORDUnet and LUMI examples.

Recommendation 5: **Federation and existing legacy should not prevent the selection of the best qualitative and cost-effective solution or to build a new one**

The federation of existing e-Infrastructure solutions and services is one of the key objectives of EOSC. It is certainly a fundamental aspect to capitalise on previous investments but it cannot be the reason to prevent new (possibly disruptive) investments that in the long term could be more profitable. One should not just accept and maintain whatever legacy comes from previous effort. The current European e-Infrastructure landscape relies almost entirely on EU funding calls, that because of their nature, are lacking long-term funding models capable of sustaining or reinforcing over time the solutions that are delivered. This situation creates a constellation of consortia, very often composed by the same set of actors, sometimes with overlapping mandates, with little or no national coordination, commitment and funding.

Little attention is paid to acquiring infrastructure and services from the best (e.g. cross-border) qualitative and cost-effective sources.

The LUMI experience shows that building something from scratch sometimes is better and more cost effective than reusing existing solutions.

Recommendation 6: **Consider the LUMI funding model to solve the EOSC computing infrastructure funding issue**

LUMI is composed of a consortium of national owners, being e-Infrastructure provider organisations, research agencies or the likes, that are willing, ready and able to contribute significant cross-border financing - assisted by EU strategic and political leadership, as well as co-financing. The value of investing in LUMI is that LUMI scales well. LUMI could be a way to generally solve the European supercomputing infrastructure needs in EOSC. Those who have already invested national funding in the consortium have secured access to supercomputing resources, while those who haven't, can still apply through competitive selection (merit) to the 50% quota paid for by the European Commission.

This model, however, doesn't work for many other types of resources such as software or network. It is suitable for a physical generic clearly defined infrastructure, with a clear offering and Total Cost of Ownership. It can only be applied to something that someone owns a share in.

However, it is important to remember that the precondition that made LUMI possible was a top-down political will (both from the EC and at national level) to implement and also to overcome local resistance towards initiatives requiring cross-border financing.

Recommendation 7: **Some share of funding in EOSC should be dedicated to the cross-border infrastructural needs of the long tail of science**

The subject specific infrastructures for the long tail of science, that have obvious cross-border scaling advantages, are currently not met by any funding mechanism, and they are completely left behind, with significant loss of research impact potential.

Subject specific infrastructures for the long tail of science cannot rely on the ESFRI/ERIC-like mechanism since they are not (and should not be expected to be) strong and internationally politically well organised.

The national e-Infrastructure provider organisations cannot support them, since they are focused on generic infrastructure and cannot invest in subject-specific e-Infrastructures without a subject and/or research group selection process or peer review allocation committee.

Serious consideration must be given to how this is resolved if open science is to be successful.

Recommendation 8: **Any new funding mechanisms identified by EOSC should come with clear rules and guidelines**

The VA funding mechanism has a good goal, i.e to promote access to many services (for the end-user) and provision of services for a wider audience (for the service providers). However, the service providers have to spend a lot of time and resources when trying to interpret and understand the rules. It is counterintuitive having guidelines that even the financial specialists have a hard time interpreting. Funding guidelines should support the application of different accounting mechanisms as much as possible, so long as the accounting systems can be audited effectively.

Recommendation 9: **The new governance framework for digital transition launched in March 2021 by the European Commission for funding multi-country projects deserves investigation by EOSC**

For funding multi-country projects, the Commission has proposed a new governance framework model (EDIC - European Digital Infrastructure Consortium)[26] that proposes a combination of investments from the EU budget, the Member States and the industry. The project investments are planned to build on the Recovery and Resilience Facility and other EU funding. The potential projects include a pan-European interconnected data processing infrastructure, the design and deployment of the next generation of low power trusted processors, or connected public administrations. This mechanism was announced while working on this document. EOSC should analyse this new framework and assess its applicability in the EOSC context.

[26] https://ec.europa.eu/commission/presscorner/detail/en/ip_21_983

Annexe 1: Template to collect information about the funding models

Name of the collaboration	
Background information	Short generic description
Type of resources provided	Please add a short description.
Procurer perspective (meaning those that are **"procuring the resources"** and NOT the ultimate end-users)	
Target procurers	
How the resources can be procured	
What triggers the demand	
How the procurers "pay" for the resources	
Main challenges in procuring resources	
Success factors	
Provider perspective	
Who owns the resources	
How the provision of resources is organised	
How the provision of resources is funded	
Cost drivers	
Main challenges in providing resources	
Success factors	
Strategic Perspective	
Strategic drivers/motivations	

Strategic or policy stakeholders	
End user perspective (end users = those that consume the resources)	
End user description:	
Main challenges	
Success factors	
Other stakeholders perspectives	

Annexe 2: Demonstrator Interview template

The interviews were conducted in a structured manner by using the following interview template. Target interviewees: EOSC Nordic demonstrators part of WP5.

- How is the technical component financed (development & long term maintenance)? Has it been created specifically for the use case?
- How is the technical component usually provided (e.g. free of charge, pay per use, et cetera)?
- Are the technical components already used across borders? If yes, can you elaborate on the provisioning/costing model?
- If the service is used only at the national level, what are the challenges and the costs of opening up the service as part of EOSC?
- Do you have any idea about potential funding and business models to sustain these services as part of EOSC?

Annexe 3: Interview results

Archaeology

Collaborating partners: Aarhus University and the Deutsche Klimarechenzentrum, DKRZ, in Hamburg.

Contact persons: Jens-Bjørn Riis Andresen, Associate professor at Aarhus University, School of Culture and Society at the Department of Archaeology and Heritage Studies at the Moesgaard Museum in Aarhus.

Senior Engineer Espen Uleberg from the University of Oslo, Database Administrator at the Museum of Cultural History in Oslo and responsible for the Norwegian database.
Anders Conrad <anders.conrad@deic.dk>

Scientific rationale: The plan is to integrate the two national databases, the Danish *Fund og Fortidsminder* and the Norwegian *Askeladden*. By integrating the respective national databases, scientists can search both countries' archaeological finds from the Viking age simultaneously.

The Vikings were widespread with widespread common material culture. So, if we find a tortoise brooch in the ground in Denmark, then we can find something similar in Norway. It makes sense from the researcher's point of view to have a common search interface.

DKZ has harvested all the data and made it available through the interdisciplinary service B2FIND and community metadata repositories. In practice, the two databases won't become one, they are each going to stay where they are, and no data is being moved. "We believe it is a risky and difficult strategy to start moving data, instead they must be housed at the responsible institution and stay there", says Jens-Bjørn Riis Andresen. This also allows the institution housing the data to decide which data should be open, and which should be closed.

Nowadays all archaeological activities are digitised. Everything from potsherd to the soil that has been dug in is photographed and everything is registered. "There is a very long tradition of systematic registration, and there are some very specific requirements for archaeological excavations and registration. Absolutely everything is digitised, and everyone perceives it as something completely natural," says Jens-Bjørn Riis Andresen. In Denmark, the national registration from excavations goes back to the 19th century, making it extra fortunate to include the Danish database.

It is necessary to develop a unique approach to each repository. It is very important to have a good overview and control over your metadata because the better you know what your metadata looks like and the better organised it is, the easier it is to get it harvested into one of these portals. Often a skilled human workforce is still required to define such specific workarounds for each scientific community.

EOSC relation: B2FIND is in the EOSC Service Catalogue

Interview with Jens-Bjørn Riis Andresen, 2021-08-12:

Archaeology research is organised according to current national borders. Historically the borders were different. This is a challenge.

The current setup is a demonstrator. If it proves to be successful, funding for a production environment will be sought, but today it is unclear where to find a financier. Denmark has country-specific challenges regarding the financing of domain-specific infrastructures, which does not apply to other European countries, hampering long-term commitments. Humanities lack a tradition of data handling like e.g. Physics.

The choice of B2FIND is due to the wish by DKRZ to develop this software, which they are doing at their own cost. The next step of cross-database search is expected during the autumn of 2021. Claudia Martens, martens@dkrz.de, is the technical contact person. B2FIND is

working OK for this use case, but there is a Horizon 2020 project ARIADNEplus, serving the archaeological community, which has chosen another tool, D4Science Infrastructure Gateway. Unfortunately, the toolchain does not align within the domain, and hopefully, there will be a common platform in the future.

There is an ongoing discussion about whether archaeological data should be regarded as administrative data or research data. Treating the data as administrative is not benevolent to research, but on the other hand, it can sometimes be a means to acquire funding and thus acquiesced.

Svenskt Hällristningsforskningsarkivs bilddatabas, https://www.shfa.se, might be a future collaborator.

Due to lack of financing, the development of the database 'Fund- og Fortidsminder' is running on the back burner. Susanna (Mahler, Enhedschef Medier?) at Slots og Kulturstyrelsen is the contact person for the operational side. This has led to the development of another database by various museums, which is not connected to Fund- og Fortidsminder.

Espen Uleberg, espen.uleberg@khm.uio.no, is the database administrator at the Museum of Cultural History in Oslo and may be able to provide information on cross-border funding models from a Norwegian perspective.

NLP Use Case

Interview with Jörg Tiedemann and Stephan Oepen, NLPL use case, 2021-08-25

Jörg and Stephan have been collaborating for a long time. Jörg's primary area is related to translation and Stephan's is semantic & opinion analysis. In the NLPL use case researchers work with different languages, but they apply overlapping technologies, techniques, and methods. The main goal is to avoid duplicative effort, lower the barrier to HPC utilisation by e,.g. MSc and doctoral students, and create an experimentation environment that gives replicability of results.

- *How is the technical component financed (development & long term maintenance)? Has it been created specifically for the use case?*

Development resources are made available by the participating Nordic research groups. Compute resources have kindly been made available by CSC and SIGMA2 as blanket accommodations. At the moment, Swedish and Danish researchers have access to the virtual laboratory under the EOSC-Nordic project.

The individual components (software and data sets) exist independently and are developed and maintained through an international open-source community. The NLPL use case is primarily about assembling many smaller pieces into a uniform and interoperable environment for large-scale computational experimentation, the NLPL Virtual Laboratory. The laboratory has

been developed specifically for and within the NLPL consortium (with much preparatory work through a NeICV-funded project that predates the EOSC-Nordic use case).

Infrastructure financing after the current project is yet to be determined. A barter scheme, where maybe Denmark and Sweden take on the responsibility of project development in return for access to computing resources, could be an option. NORDFORSK has issued calls for community networks that could potentially be a source of financing. The technical know-how and staffing are available within the research groups. The most promising idea is probably LUMI, as seen below.

- *How is the technical component (the virtual language laboratory) usually provided (e.g. free of charge, pay per use, et cetera)?*

NLPL primarily uses IAAS, with no higher-level services from the EOSC catalogue. During the development projects, researchers from other countries are welcome to use the resources provided by CSC and SIGMA2, but when the project ends, researchers from other countries are likely to be thrown out. It would be possible to set up an instance of the virtual lab in each country, but there are clear advantages by concentrating the resources, in addition to the apparent extra work to duplicate the effort in additional locations, not the least that researchers can cohesively see all the results in one place.

Currently, the access to the use case is through allocation and access management on the two national superclusters that host the virtual laboratory, i.e. access mechanisms managed by CSC and Sigma2 in Finland and Norway, respectively. University research groups from all consortium members gain no-cost access and some allocations of compute hours by virtue of being part of the EOSC-Nordic use case.

- *Are the technical components (the virtual language laboratory) already used across borders? If yes, can you elaborate on the provisioning/costing model?*

The virtual language lab maintains packages as part of the lab to ensure uniformity across systems. Provisioning an environment for replicable research is important using clean data sets available in the lab. The bespoke EasyBuild recipes are not unique to NLPL and could potentially be of interest to others.

Yes, there are currently two parallel instances of the NLPL virtual laboratory, one in Finland, one in Norway (on Puhti and Saga today, Taito and Abel in the past, Betzy is likely to be added); both are accessible to all consortium members, and both see substantive use from outside their home countries.

- *If the service is used only at the national level, what are the challenges and the costs of opening up the service as part of EOSC?*

The service is used already across borders.

- *Do you have any idea about potential funding and business models to sustain these services as part of EOSC?*

The long-term vision is that national e-infrastructures should pool resources and fund this together to enable collaboration through better sustainable infrastructure. The Dellingr project looked into this, but unfortunately, it did not come up with conclusive solutions. There is also a need for moderate funding to help the community to arrange itself.

Language processing is not a traditional HPC discipline, and its way of working is different.

That is the million-dollar question :-). Possibly LUMI can provide a future home for the virtual laboratory (once we can get enough of the software working well there, and provided that NLP research will be considered in-scope for LUMI). Then at least current NLPL members and additional collaborators from, e.g., the Czech Republic and Switzerland could hope to gain access and allocations through national mechanisms, i.e. get to the same, shared meeting place each through their entrance door. This solution would combine a shared environment where all the researchers can see all results and simultaneously provide allocations from their national partitions of LUMI.

NLPL is already a pilot user of LUMI, using the Norwegian and Finnish partitions.

FAIR Climate Data for the Nordics

Collaborating partners:

- Linköping University, NSC, Sweden
- Danish Meteorological Institute (DMI), Denmark
- IT Center for Science (CSC), Finland
- NORCE, Norway
- The Norwegian Meteorological Institute (MetNo), Norway
- University of Iceland (UoI), Iceland

Contact persons:

- Hamish Struthers <struthers@nsc.liu.se>, LiU/NSC, Sweden
- Anne Claire Mireille Fouilloux <annefou@geo.uio.no>, UiO/Dept. of Geosciences

Scientific rationale:

The goal is to provide FAIR climate data to all Nordic communities interested in climate mitigation and climate change impact assessment. This use case aims at serving NeIC NICEST-2 (the second phase of the Nordic Collaboration on e-Infrastructures for Earth System),

using Galaxy (portal) HPC in Norway and cloud in Sweden and Finland. S3-accessible storage and B2Drop (at CSC)

EOSC relation: B2DROP and B2SAFE are the EOSC Service Catalogue

Interview with HS, NICEST use case, 210819

- *How is the technical component financed (development & long term maintenance)? Has it been created specifically for the use case?*

The galaxy portal utilised by the NICEST use case is “piggybacked” on the ELIXIR portal hosted in Norway. While there was originally supposed to be a galaxy portal for EOSC-Nordic, it was in the end not deemed viable.

The galaxy portal was chosen due to availability - it was available in one institution within the community. For the time being, it is provided as-is at no cost. Financing after the project hasn’t been discussed.

- *How is the technical component (the Galaxy portal) usually provided (e.g free of charge, pay per use, etc)?*

NICEST doesn’t have any funding for buying services, licenses, or similar, there are only human resources available. Thus services available for free, primarily provided by service providers that can be provided in kind, have been chosen. There is a focus on technical details, and financial aspects are not areas of focus.

Sustainability is not a focus in NeIC projects, or within EOSC-Nordic WP 3. The cost for a production phase, which will probably be much longer, may be affected by the choices made during the setup, but this is beyond the scope of current work.

No criteria have been set up, from which services could have been chosen. Instead, Galaxy has been used, primarily because there had been some previous work done by Oslo university using Galaxy, and there was a developer familiar with the instance down the hallway relative to the personnel in the project. Galaxy was available and chosen due to availability.

- *Are the technical components (the Galaxy portal) already used cross border? If yes, can you elaborate on the provisioning/costing model?*

The NICEST use case uses cloud services within EOSC-Nordic. These services are made available by the participating national providers, such as SNIC and CSC. There is currently no cross-border provisioning model beyond the pilot, and it hasn’t been discussed.

In-kind style of thinking drives the choices of the project, which may be well for pilots and similar, but may not work or be the best possible solution in a production environment. For production, it

may make sense to buy services, commercial programmes, or both to obtain a defined level of expectation and cap maintenance costs.

- *If the service is used only at a national level, what are the challenges and the costs of opening up the service as part of EOSC?*

Technically it would be trivial to get the galaxy portal to operate as a service in the EOSC catalogue. However, establishing it as a service with a business model, defined service level, etc. is nontrivial.

Within the framework of WP 5 deliverables, the focus is on technical solutions rather than issues like business models. There has been no discussion about costs.

Sharing research data is one aim of the project. How this will be implemented will depend upon the kind of back-end data services used. The data lake, which is part of Elixir efforts, provided by CSC seems promising for data sharing and exploits the existing systems. The intent is that there should be no associated costs for the data providers. Depending on the technical solution, costs might be incurred.

Data management is not a priority for researchers, and it is difficult to motivate them to do anything beyond the minimal effort required by funders. On the other hand, there are additional costs associated with the reuse of raw data, as it requires lots of time and effort to be made usable by other researchers.

- *Do you have any idea about potential funding and business models to sustain these services as part of EOSC?*

No idea. Projects would probably benefit from an increased focus on economic and legal issues. The goal for most projects is to go into production.

Precision Medicine

Medication in Pregnancy & Childhood Cancer

Collaborating partners:

- Swedish National Infrastructure for Computing (SNIC)
- IT Center for Science (CSC)
- Computerome (Technical University of Denmark)
- University Center for Information Technology at University of Oslo

Contact persons:

- Gard Olav Sundby Thomassen, University of Oslo (UiO)?

Scientific rationale:

The study is a population-based registry linkage study aiming to investigate the association between maternal use of antibiotics in pregnancy and the risk of childhood cancer. The ultimate goal is to determine which antibiotics are safe to use in pregnancy regarding the future risk of cancer in the child.

The study will use data on pregnancy complications, redeemed prescriptions, and cancer diagnoses, all legally considered sensitive data. Childhood cancer being a very infrequent outcome, it is necessary to combine data from health registries in Denmark, Finland, Norway, and Sweden. The challenge is to prepare the legal and technical risk assessment to actually move data across borders. The plan is to store the data from different sources in Norway at the UiO-operated *Tjenster for Sensitive Data,* TSD (Services for Sensitive Data), an infrastructure especially made for such purposes. The TSD is a platform for collecting, storing, analysing, and sharing sensitive data complying with the Norwegian privacy regulation. The TSD is primarily an IT platform for research, but it has also been used for clinical and commercial research. TSD is developed and operated by UiO.

Data comes from health registries of approximately 2 million children and their parents. The project aims at sharing data on a single platform, allowing the data to be analysed in a combined manner, as opposed to other models in which data is analysed separately in each country and combined through meta-analysis. The idea is to make data accessible through a secure and encrypted VPN, enabling the PI in Norway and collaborators in Denmark, Finland and Sweden to analyse data from all four countries as one data file. The plan is to use the Research Infrastructure platform TSD as a testbed for sending data from Denmark to Norway.

EOSC relation: none?