CSC-IT Center for Science

Project Charter

Version 1.9.

EOSC Federation: EOSC Finland Node

1. Project Summary

CSC - IT Center for Science is contributing to the EOSC Federation by building up a national node for Finland. The node will be built in collaboration with the Finnish Ministry of Education and Culture, the national Open Science Coordination and other national stakeholders that will be gradually engaged (and onboarded in the national node) via the EOSC Finnish Forum mechanism[1].

The EOSC Finland Node will unlock access to national FAIR data, data services and research data management competence development. These resources will be accessible via a user-friendly entry point. The node will be accessible for Finnish researchers via the Haka federation, and it will be integrated with MyAccessID and EOSC AAI. The datasets, data services and training resources will be discoverable via the EOSC federated service catalogue. Through coordination with the LUMI AI Factory the national EOSC node will provide access to unique and large datasets and will also pilot a cross-node use cases on secure processing environments. As CSC is the main contractor of the EuroHPC Federation (EuroFP) tender, CSC will also promote interoperability between the EOSC Federation and the EuroFP.

2. Value Proposition

Main Goals

- making Finnish FAIR datasets, services and RDM trainings discoverable in the EOSC Federation (examples of national services are: fairdata.fi; research.fi; national data infrastructures and RI national nodes (e.g. FinCLARIN); etc.)
- making available unique, large size or restricted access datasets via interoperability with the LUMI AI factory
- making sensitive data discoverable and through authorisation accessible in a secure processing environment (for example in collaboration with EUDAT, ELIXIR, SURF)
- making the EOSC Finland node interoperable with EuroFP and the LUMI AI factory
- providing all the necessary core federating capabilities towards the EOSC Federation
- contributing to the interoperability framework and joint concept model development

Needs addressed

- need for FAIR supporting user-centric services that facilitate the research process
- need for RDM competence development
- availability of unique, large size and restricted datasets
- integrating federating capabilities

Key Benefits

- demonstrating the value of a national node for the EOSC Federation as a whole and for the country
- increasing discovery and availability of FAIR data and FAIR supporting services
- testing interoperability between EuroFP and EOSC Federation

[1] EOSC Forum Finland https://avointiede.fi/en/networks/eosc/eosc-finnish-forum

- bringing the EOSC Federation close to the private sector and their RDI collaboration with research performing organisations thanks to the connection with the LUMI AI factory

Who Benefits

- European researchers who can access Finnish datasets, services and RDM training resources on top of the other resources made available through the EOSC Federation
- European researchers and businesses that can access large size or restricted datasets
- European researchers who need to access and process sensitive data
- Finnish organisations willing to contribute to the EOSC Federation via a national node
- Stakeholders that will benefit from the integration of the EOSC Finland node in the LUMI AI factory ecosystem

3. USE CASES

Focus will in the first phase be put on use cases EOSCFI-UC-1, EOSCFI-UC-2 and EOSCFI-UC-4.

Use Case ID	Use Case Description	Federation Contributions & Value to Users
EOSCFI-UC-1	Datasets-as-a-Service will be the primary data offering of the LUMI AI Factory based in Finland. Connecting the AI factory to the EOSC Finland Node will provide the EOSC Federation datasets of exceptional value, typically not available to users due to their large size or restricted access. Additionally, it will provide a good selection of typical AI datasets. The plan is to make these datasets/services discoverable and available in the EOSC Federation. Datasets with link to scientific multi-node use case will be prioritised. • Developing and testing the concept and productisation of datasets-as-a-service • Creating data contracts and developing the governance model • Making datasets visible in catalogs	*Federation contribution: discoverability and availability of previously not accessible major datasets via the EOSC Federation* *Value to users: access to previously not accessible datasets*
EOSCFI-UC-2	Making sensitive data discoverable and accessible through authorisation in secure processing environments in collaboration with EUDAT, ELIXIR, SURF; • Making sensitive datasets findable across nodes • Looking at AAI solutions (with output from EOSC ENTRUST) • Enabling access and multi-node workflows	*Federation contribution: multi-node workflow* *Value to users: access to sensitive data in a secure processing environment*

EOSCFI-UC-3	Onboarding of resources in the Finnish node.	*Federation contribution: making Finnish resources discoverable in the EOSC Federation increasing their reuse* *Value to users: access to FAIR data and services and RDM training resources*
EOSCFI-UC-4	Integration with EU Node (AAI and federated catalogues)	*Federation contribution: enabling data and services discoverability* *Value to users: multi-node interoperability*
EOSCFI-UC-5	Ensuring sustainable EOSC collaboration of the CSC RDM Competence Center	*Federation contribution: Continue building the pan-European network of competence centres to speed up the training of European researchers and harmonise the training of new professional figures for scientific data management set up within Skills4EOSC.* *Value for users: stronger research data management competences and support*

In-scope for the EOSC Finland Node and use cases

- Entry point for the EOSC Finland node based on MyCSC customer portal technology and end-user processes
    - MyAccessID integration for EOSC users, with well-defined set of CSC services and their accesses
- Core functions provided via the node: single-sing-on AAI integration, federated service catalogue, helpdesk for users, monitoring, usage reporting
- Participation in development of federating capabilities for the EOSC Federation: AAI, resource allocation and credit concept model for services and resources
    - CSC resource allocation concept model can be shared as a reference: services, user project governance model, accounting ("billing units"), allocations, usage reporting
- Stakeholder support: European researchers, Finnish researchers, Finnish organisations willing to onboard resources in the EOSC Finland Node (e.g. Finnish universities, national data infrastructures, RI national nodes, etc)
- Integration:
    - Integration with EOSC EU Node AAI and federated catalogue
    - Pre-study of possible integration with ELIXIR, EUDAT and SURF nodes for the multi-node workflow around sensitive data
    - Integration with new potential nodes for other waves (e.g. CLARIN)

Out of scope for the EOSC Finland Node and use cases

- The EOSC Finland Node will sustain in production only the resources and workflows that have proven to bring added value to the stakeholders. Experiments

that will be done during the build up phase that do not generate impact for Finland or for the EOSC Federation will be dismissed.
- Most services will be provided in the first phase as-is, for example with their current access and usage policies as set by the funders that may have restrictions in scope of support, geographic reach or target audiences.

4. External Dependencies & Key Risks

The dependencies are described more specifically per use case below in chapter 6 (pp. 7-8).

External Dependencies & Risks	Actions	Deadline
Different timelines, maturity levels and resources available from other nodes, incl EOSC EU Node.	Discussing the Finland node timeline with the other related actors	May 2025

5. Contributions

Access policies and restrictions for the resources of the Finnish node

Services and data:

- **Existing access and use policies for services using quota or billing units require a Project Manager affiliation with a Finnish RPO approved by the Ministry of Culture and Education.**
- **Access to datasets or metadata is not restricted based on geography or affiliation.** Overall open access policies for research data are widely mandated and restrictions exist only where necessary.
- More detailed policies:
  - https://research.csc.fi/terms-of-use/
  - https://research.csc.fi/free-of-charge-use/

- **Services onboarded to the Finnish node have their respective policies,** e.g. the Language Bank of Finland https://www.kielipankki.fi/language-bank/terms-of-use/, Geodata download service https://paituli.csc.fi/home.html
- Only qualified services with clear policies, terms of use and service description, will be onboarded according to the principles of the EOSC Federation
- Regarding datasets, see the policies for CSC services for research, including the data management policy and PID policy https://research.csc.fi/policies/
- Service agreements will be done separately for each Dataset-as-a-Service, as open as possible and as restricted as necessary.

Data catalog (Metax):

- User interface https://etsin.fairdata.fi/
- Data CC0 (including reference datasets)

Research information (research.fi):

- User interface https://research.fi/

- Needs contract but is free. OpenAIRE has an agreement and access to this dataset.

Service catalog (data.csc.fi):

- User interface https://research.csc.fi/service/
- CC0

Deliverables

ID	Deliverable Name	Deliverable Description	Deliverable Owner
1	Project charter	Description of the purpose, value proposition and use cases of the EOSC Finland Node	CSC

2	Analysis of the EOSC EU node federating capabilities and plan for integration	Analysis of the EU Node federating capabilities and plans for integration. Necessary specifications for federation are
		
		Mappings and crosswalks:
		• Service catalog (data.csc.fi, open data)
		  ○ EOSC FI > EU Node
		  ○ EU Node > EOSC FI
		  ○ Possibly other nodes with machine-readable catalogs and relevant services or resources
		• Datasets (https://www.fairdata.fi/en/metax/, open data)
		  ○ EOSC FI > EU Node
		• Suggestion for presenting Datasets-as-a-Service as federated resources
		• Other CRIS information if relevant
		  ○ (https://research.fi/en/service-info, needs contract, free)
		  ○ EOSC FI > EU Node
		
		Architecture and requirements:
		• AARC Blueprint Architecture
		• Analysis of access policies
		• AAI solution and security
		  ○ Analysis of the different AAI systems in use, including eIDAS
		  ○ Taking part in discussions within the federation regarding specifications, levels of trust, security architecture etc.
		  ○ Identifying possibly needed adjustments in existing solutions
			CSC

3	Service documentation	EOSC Finland Node IT Governance documentation Architecture Design Plan and the Architecture Canvas: including hosting, DNS, network	CSC

		IT Security Plan: security model, security architecture and IT security impact assessment Technical Data Protection Plan and Data Protection Impact Assessment (questionnaire) Operational Disaster Recovery Plan (including implementation plan) Evaluation of the IT Security Plan: provide input to the IT Security Risk Report Service Interoperability Plan Node Policies	
4	Stakeholder and community engagement strategy	Analysis of stakeholder groups and a communication and engagement plan to reach and engage relevant stakeholders. Relevant guidelines and documents for onboarding and interacting with the Finnish node, its services and resources for different target groups.	CSC
5	Multi-node use case	Implementation plan for multi-node use case, possibly with restricted data	CSC

6. Timing and Milestones

- **Start Date:** 17-03-2025.
- **Expected Duration:** 2 years organised in phases; 1st Phase 17.3.2025 – 1.1.2026.
- **Productization** of the node will be prepared in the first phase and done in the beginning of the second phase after assessment.
- A separate plan for the 2nd phase will be made based on assessment of the pilot node, the EOSC federation and the use cases, as well as the development in related projects, such as EuroFP, LUMI AI Factory and EOSC-ENTRUST.

Key Milestones for 1st phase 17.3.2025 – 1.1.2026

ID	Milestone Description	Target Delivery Date
1	First specifications for the pilot node ○ Relevant mappings and API information to be shared in a first version of Deliverable #2 to enable easy federation Stakeholder engagement plan v.1	April 2025
2	Preliminary description of at least one multi-node use case	May 2025
3	First concept and requirements for the Dataset-as-a-Service ○ Suggestion for publication process	June 2025
4	First instance of the EOSC Finland node	August 2025

ID	Milestone Description	Target Delivery Date
5	EOSC EU Node AAI and federated catalogue integration ○ Critical dependency to collaboration with and development in the EU Node (May – September)	September 2025
6	Datasets-as-a-service pilot ○ Dependency: enabling publication in a federated catalog requires interoperability with receiving node	October 2025
7	Demonstration of multi-node use case ○ Critical dependency to other nodes (May – September)	October 2025

7. CONTACT & SUBMISSION

Espoo, Finland 31.3.2025.

Role	Name	Email
Coordinator	Jessica Parland-von Essen	jessica.parland-vonessen@csc.fi
Operation Manager	tbd	
Security Officer	Urpo Kaila	security@csc.fi ; urpo.kaila@csc.fi
Scientific Officer	Martin Matthiesen	martin.matthiesen@csc.fi
Legal/privacy officer	TBD	privacy@csc.fi until specified