

CERN

Project Charter

CERN EOSC Node

June 2026

1. Project Summary

During the 2025 EOSC Federation Build-up Phase, CERN participated as a first-wave candidate EOSC Node with the objective of validating interoperability between CERN services and the emerging EOSC Federation. This work focused on three core dimensions: metadata integration and discoverability, federated access via EOSC AAI, and deployment of scientific workflow and data services on EOSC infrastructure. CERN successfully onboarded mature production services and technologies including Zenodo, REANA, Indico, CERN Open Data, INSPIRE-HEP, CERNBox and Virtual Research Environment (VRE)-related capabilities, while contributing to cross-node scientific workflows and EOSC-level technical subgroups. Key achievements included exposing more than 5 PB of CERN collision data, related simulations, and HEP research outputs through EOSC discovery mechanisms, enabling federated access to selected services through EOSC AAI, and demonstrating distributed scientific workflows across multiple EOSC Nodes using REANA. The experience and operational lessons gained in 2025 established the technical and organisational foundations required for the next phase: transitioning from interoperability pilots towards a stable and operational CERN EOSC Node.

In 2026, CERN will deliver the first version of the CERN EOSC Node MVP, targeted for October 2026. The MVP will be a lightweight EOSC Node that will showcase the following capabilities: an Authentication and Authorization Infrastructure (AAI) compliant with EOSC requirements, a catalogue of resources and services that CERN provides or plans to provide through the EOSC Federation, a helpdesk for user support, and a monitoring framework for services, together with a Service Management System compliant with the FitSM standards.

The services exposed through the CERN Node will include, among others, Zenodo/InvenioRDM, the CERN Open Data, Indico, CERNBox, distributed data access and management via Rucio, REANA and VRE for data analysis, INSPIREHEP and HEPData, and related training material.

The development roadmap will follow an incremental vision: first, a practical MVP and education/outreach-oriented entry point; second, an education and research node for open data; and third, a multi-tiered research node connected to CERN infrastructure. The second and third stages will depend on governance decisions, resource availability and, in particular, EOSC MESH funding for implementing federating capabilities.

The CERN Node will foster the EOSC Federation by providing access to a catalogue of mature open-source technologies and research outputs, available via federated access; moreover, it will allow the discovery and usage of high-value FAIR HEP data, and the possibility of developing new use cases and new analyses. The scope is to support not only CERN researchers and educators, but notably the external collaborators and the EOSC Federation community as a whole. Moreover, the development of new use cases, and the operational experience gained by co-designing the federation with the other EOSC Nodes is a focal point of CERN commitment.

2. Value Proposition

The **main goal** is to establish a CERN EOSC Node that provides a trusted, visible and operational entry point to CERN open science services, HEP open data, reusable computational research technologies, particle physics analyses, and federated scientific workflows.

The project addresses the need for:

- a clear EOSC entry point for CERN open science infrastructure, services and technologies
- better discovery of CERN and HEP research outputs through EOSC catalogues
- federated access to onboarded CERN services through EOSC AAI
- reusable analysis workflows and reproducible science environments
- cross-node scientific workflows using CERN technologies such as REANA and Rucio

- structured support, monitoring and service management for CERN's EOSC-facing services
- a pathway enabling CERN services and infrastructure to participate in the EOSC Federation in a coordinated and consistent way, avoiding duplication of existing CERN platforms and ensuring strategic alignment across CERN's internal projects and activities
- a standardised way to bridge the interaction between the EOSC Federation and other distributed computing infrastructures, such as WLCG

Key Benefits:

- For EOSC, CERN contributes mature, production-grade services and technologies already used by large scientific communities. CERN also contributes expertise in data preservation, reproducibility, workflow execution, distributed computing, open-source repository platforms, scientific collaboration tools and exabyte-scale data management
- For CERN, the Node provides a structured framework to align CERN's Open Science strategy with EOSC, improve interoperability of CERN services, increase the impact of CERN technologies, and create a path for CERN researchers to benefit from federated European research infrastructures
- For user communities, the Node provides easier discovery and reuse of CERN open data, scientific software, training material, workflows and analysis platforms. It also creates a potential route for trusted external researchers, such as theorists and phenomenologists, to analyse LHC (and previous accelerators) open data close to where the data resides

Who Benefits:

- EOSC Federation and EOSC Nodes
- Researchers using CERN and HEP open data
- Educators and training communities
- HEP theorists and phenomenologists without direct access to CERN infrastructure
- CERN researchers, once the Node evolves into a feature-full research platform simplifying cooperation with external researchers
- Open Science policy, infrastructure and interoperability communities

Roadmap and Future Vision

Stage 1: 2026 CERN EOSC Node MVP, targeted until October 2026

The first implementation stage will deliver the CERN EOSC Node MVP. This will be a lightweight, realistic and operational first version of the Node.

The MVP will provide:

- a live CERN EOSC Node website
- EOSC AAI integrated
- helpdesk/support routing
- service monitoring, and operational contact points
- a catalogue of all CERN services provided or planned for EOSC
- clear descriptions of reusable CERN technologies available to other nodes
- links to CERN Open Data, Zenodo, INSPIRE-HEP, HEPData and relevant repositories
- basic access policy information
- documentation for users and EOSC partners
- curated examples of analysis and teaching workflows

The purpose of the CERN Node is to better support the scientific use of LHC data and related services by CERN communities and external contributors participating in the Federation. The MVP **will not** provide unrestricted compute resources or support large-scale production workloads. Instead, its purpose is to establish visibility, clarity, federation compliance, and an operational foundation for future growth in support of these scientific activities.

Stage 2: Education and Research Node for Open Data

After the MVP, the CERN Node should evolve towards an education and research node for CERN open data. This stage will build on the MVP and provide more structured support for analysis, training, reproducibility and trusted research use cases.

This stage will particularly depend on EOSC MESH funding (starting in September 2026) for implementing federating capabilities, especially where CERN needs to integrate with federation-level services such as AAI, catalogues, helpdesk, service monitoring, service management, and federated analysis workflow execution.

The focus will be on:

- controlled compute access for vetted workflows
- mechanisms for trusted researchers to request additional quotas or capabilities
- collaboration with second-wave EOSC Nodes on agreed cross-node use cases

Stage 3: Multi-tiered Research Node with CERN Integration

The third stage is fundamental because the CERN Node must not remain only an external-facing outreach or an “enhanced” open-data portal. To be sustainable and valuable for the HEP community, it must become well integrated with CERN infrastructure and support CERN researchers as well.

Stage 3 will evolve the Node into a multi-tiered research-capable platform within the EOSC Federation. CERN users would be able to access selected baseline services through the Node, while external users would access services according to defined policies, quotas and trust levels.

This stage is important because it would:

- align EOSC participation with CERN’s internal infrastructure strategy
- avoid duplication between EOSC-facing and CERN-internal services
- provide tangible value to CERN researchers
- allow CERN users and non-CERN users to seamlessly collaborate in the same research platform
- allow CERN users and experiments to access selected EOSC or external resources where useful

3. Use Cases

Use Case ID	Use Case Description	Federation Contributions & Value to Users
-------------	----------------------	---

UC-1	Vetted open data analyses	Allow any EOSC user to run vetted analyses over 5 PB of CERN Open Data for education and outreach purposes.
UC-2	Federated scientific workflow execution	Allow research workflows from collaborating remote EOSC Nodes to perform open data analyses close to where the data sits. Support interoperability between compute, workflow and data services across EOSC Nodes and help validation of distributed scientific analysis paradigm. (A part of the EOSC Mesh project.)
UC-3	Reproducibility checker	A researcher performs statistical analysis and would like to verify its computational reproducibility beyond the original computing environment. The researcher (or the EOSC Node A infrastructure) will package the analysis code, the container environment, the input data and parameters into an “encapsulated analysis tarball” and ship the said analysis capsule to another EOSC Node B to verify the results and inform back. (A part of the EOSC Mesh project.)
UC-4	LHC Open Event Generation data analysis	Collaboration between the CERN Node and LHC BSM Reinterpretation working group, to leverage the ATLAS / CMS Open Event generation and enable researchers that do not have access to CERN internal resources to use the data and the computing power of the federation. For more information about use cases and scope, please see Sections 3 and 5.3 of arXiv:2605.12229v1
UC-5	CMS open data processing and simulations	Allow researchers to run open data processing and simulation workflows (such as AOD, MiniAOD and NanoAOD data processing) across the EOSC Federation Nodes. Helps to test the open data reproducibility recipes and to benchmark the EOSC Federation distributed compute model.
UC-6	LHCb open data simulation production	Allow researchers to run LHCb open data simulations on EOSC Federation Nodes, complementing LHCb collision data ntuples with simulation capabilities essential for sophisticated physics analyses. Helps to benchmark the EOSC Federation distributed compute model.
UC-7	ATLAS RECAST	Allow researchers to trigger running of parametrised open data reinterpretations without having access to neither original data nor original code. Helps researchers to test novel theoretical ideas against past data analyses. Offers EOSC Federation Nodes a use case of offering compute services over

		"closed" data. (A similar use case may exist in Metrofood-RI food and nutrition data services as part of the EOSC Mesh project.)
UC-8	EOSC Gravity grant: Distributed Data Management between HEP and Genomics	Collaboration between CERN Node, the Italian Node, and the Area Science Park research infrastructure about federating genomics data through Rucio

4. External Dependencies & Key Risks

External Dependencies: For the initial stages, the CERN Node is quite independent but it will need to follow the recommendations coming out of the NCC and NOC.

Key Risks & Mitigation Measures:

Risk / Dependency	Impact	Mitigation
Scope creep from external compute requests	Demand may exceed available CERN resources.	Limit MVP to vetted workflows; provide minimal quotas.
Security risks from external workloads	Arbitrary workloads could create operational or cybersecurity risks.	During stage 1, restrict execution, require vetting, sandboxing, monitoring and trust-based access.
Overlap with existing CERN infrastructure	CERN users may see limited value if EOSC services duplicate existing tools.	Use Stage 3 to integrate with CERN infrastructure and complement internal services.
Sustainability and operational burden	Support, monitoring and service operation may require additional resources.	If the Federation rules and procedures impose a significant operational burden, they should be reviewed and streamlined to enable nodes to comply with a core set of requirements while retaining sufficient flexibility in their implementation.
Governance and access policy complexity	External/CERN user models may become difficult to manage.	Define access tiers and policies before expanding beyond MVP.
Cross-node dependencies	Use cases depend on readiness of partner nodes.	Keep use cases modular; document dependencies; avoid making MVP dependent on external readiness.

5. Contributions

- **Role in Project:**
 - CERN acts as a first-wave EOSC Node contributor, service provider, technology provider, open data provider and scientific workflow infrastructure contributor.
- **Main Contributions:**
 - CERN EOSC Node MVP
 - service/resource descriptions
 - selected service catalogue entries
 - AAI integration experience
 - open data and repository resources
 - REANA reusable analysis workflow technology
 - VRE/Jupyter-based analysis environments
 - knowledge transfer around InvenioRDM, Zenodo, Indico, CERNBox and Rucio
 - participation in EOSC Mesh and cross-node workflow use cases
 - operational experience from production-grade scientific services
- **Deliverables:**

ID	Deliverable	Description	Owner	Status
D1	CERN EOSC Node MVP	Initial CERN Node describing services, resources, and technologies, as well as an initial access policy and resource allocation policy.	CERN	Planned for October 2026
D2	Analysis catalogue	Curated list of research analyses that can be run by EOSC-authenticated users.	CERN	Planned for October 2026
D3	Federating capabilities implementation plan	Plan covering AAI, catalogues, helpdesk, monitoring and SMS.	CERN	Sept 2026
D4	Roadmap for Stage 2 and Stage 3	Post-MVP plan, including EOSC MESH-funded activities.	CERN	Q4 2026
D5	REANA cross-node workflow demonstrator	Reproducible workflow use case with cross-node relevance.	CERN / partners	Ongoing
D6	Support and operations model	Basic support, escalation, service ownership and operational contacts.	CERN	Q2 2027
D7	Use-case implementation plan	Updated use-case list aligned with second-wave collaborations.	CERN	Q2 2027
D8	Roadmap for Stage 3	CERN integration path.	CERN	Q4 2027

6. Timing and Milestones

- **Start Date:** June 2026
- **Expected Duration:** 6 months for Stage 1
- **Key Milestones:**

ID	Milestone Description	Target Delivery Date
M1	CERN MVP passing the EOSC AAI validation test	Aug 2026
M2	Resource inventory completed: Services, repositories, technologies and access policies reviewed.	Sept 2026
M3	Federating capabilities plan agreed: AAI, catalogues, helpdesk, monitoring and SMS plan validated.	Sept 2026
M4	CERN EOSC Node website alpha: Internal draft of MVP website available for review.	Sept 2026
M5	Use cases confirmed: Active CERN-relevant use cases confirmed; abandoned/non-relevant use cases removed.	Q3 2026
M6	MVP launch: First version of CERN EOSC Node live website released.	Oct 2026
M7	Stage 2 implementation planning: Education and Research Open Data Node plan prepared, including EOSC MESH-funded activities.	Nov 2026
M8	Federating capabilities implementation: Implementation of enhanced AAI, catalogues, helpdesk, monitoring and SMS.	2027 onwards
M9	Stage 3 assessment: Assessment of CERN infrastructure integration and value for CERN researchers.	2027 onwards

7. Contact & Submission

Role	Name	Email
Coordinator	Jose Benito Gonzalez Lopez	jose.benito.gonzalez@cern.ch
Deputy Coordinator	Tibor Simko	tibor.simko@cern.ch
Operation Manager	Hannah Short	hannah.short@cern.ch
Deputy Operation Manager	Giovanni Guerrieri	giovanni.guerrieri@cern.ch
Security Officer	Stefan Lüders	stefan.lueders@cern.ch,

Scientific Officer	Micha Moskovic and Xavier Espinal	micha.moshe.moskovic@cern.ch, xavier.espinal@cern.ch
Legal/Privacy Officer	Lea Guglielmetto-Chaleard, PMO in consultation with CERN's legal service	lea.guglielmetto-chaleard@cern.c h
Communication Officer	Antonella Del Rosso	Antonella.Del.Rosso@cern.ch