

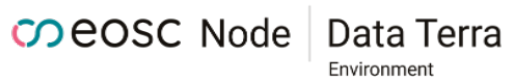
CNRS

Project Charter

Updated Version (V2.1)

6 July 2026

EOSC Node Data Terra



1. PROJECT SUMMARY

Data Terra is the French national Research E-Infrastructure for Earth system, environment and biodiversity sciences. It federates five domain data hubs, including the French thematic contributions to European research infrastructures (ESFRI/ERICs), and develops cross-cutting services and expertise, with a particular focus on satellite observations. As a key component of the French research infrastructure landscape, Data Terra also contributes to major European initiatives such as Copernicus by providing data and services. Recognized on the French national Roadmap of Research Infrastructures, it is considered strategically important both nationally and as a major contribution to the European Research Area (ERA).

Data Terra supports a cross-domain approach to address major societal transitions and the Sustainable Development Goals (SDGs). Its strategy is built around four priorities: advancing research in line with emerging scientific challenges; providing long-term, high-quality FAIR multi-domain data and interoperable metadata; strengthening the national research ecosystem through shared data services and infrastructure access; and reinforcing collaboration with European and international partners.

Data Terra is also the French thematic reference centre for Earth, Climate, Environment and Biodiversity within Research Data Gouv, reflecting its leading role in open science and the development of international data standards. In addition, it is a partner of AI Factory France, led by GENCI, where it coordinates the Earth sciences domain and contributes to strengthening links between AI research, industry, and environmental sciences.

During the 2025 EOSC Federation build-up phase, the EOSC Node Data Terra, one of the thematic nodes representing the Earth system and environmental sciences within the EOSC Federation, launched an initiative to expand its node by involving partners and contributors from other EU Member States. The objective is to establish a European transnational thematic node. The first step in this expansion is the signature of a Memorandum of Understanding (MoU) between Data Terra and the German research infrastructure **NFDI4Earth**. Under this MoU, which is currently being signed, Data Terra and NFDI4Earth aim to establish common approaches and frameworks for service integration, interoperability, and quality assurance, including the implementation of FAIR principles, technical standards, and the consideration of ethical and legal aspects.

Data Terra and **NFDI4Earth** share indeed a common mission to address the digital needs of the Earth system science (ESS) community and enable cutting-edge research. Both infrastructures aim to consolidate and harmonize research data services, providing researchers with a user-centric, innovation-driven, efficient, and sustainable service ecosystem at the national, European, and global levels.

1.1 Status of adoption of federating capabilities

In 2026 Data Terra will deliver the first version of the EOSC Node with the initial operational release planned for October 2026.

The Data Terra **Node's SSO/AAI system** has been in production since February and connected to MyAccessId since early March. It complies with the technical requirements (AARC blueprint architecture specifications) and regulatory requirements (privacy notice, incident response, etc.) defined by the EOSC AAI WG. The services that will be exposed towards the Federation since the early beginning of the operational phase have been connected (or are in the process of being connected). Additional services identified in the catalogue requiring authentication are expected to be connected by the end of 2026.

Concerning the **Node's Resources Catalogue**, Data Terra has adopted a foundational design choice based on a unified catalogue model built on DCAT. In this approach, every resource is exposed as a dcat:Resource, while its specific nature is indicated via the dcterms:type property.

Datasets are additionally typed as `dcat:Dataset` to ensure full DCAT-AP compliance, whereas other resource types (e.g. Software, Publications, Services, etc.) are represented solely through `dcterms:type`, linked to a controlled SKOS vocabulary. Data harvesting is performed through a dedicated component provided by the French Gaia Data project (MTEP). In parallel, specific connectors are being developed to aggregate additional categories of research outputs, including CodeMeta for software, DataCite for publications, and GeoDCAT-AP for geospatial data and services. A key advancement is the design of an intelligent EOSC connector acting as a local quality gate. Before any metadata is exposed to the EOSC EU Node (EEN), it undergoes strict validation ensuring the presence of mandatory Persistent Identifiers (PIDs), licenses, and core metadata fields. This mechanism guarantees a “zero rejection” policy by the European Knowledge Graph. The system is designed to preserve rich, domain-specific metadata locally (e.g. spatial and temporal extent, fine-grained provenance) while projecting only a strictly compliant subset to the EEN. Finally, metadata records are enriched through a combination of rule-based processing and AI-assisted enrichment techniques. The final version of the resource catalogue is expected to be available by the beginning of October 2026.

For the **Helpdesk** and **Monitoring** functions, the EOSC Node Data Terra currently relies on distributed infrastructures operated by the node’s data centres and service providers at both the national and European levels. To support the federation of these services across the EOSC Node Data Terra, discussions have been initiated with EGI in the context of its role in the EOSC EU Node, as well as within the EOSC Beyond and EOSC Mesh projects, in which Data Terra participates. The objective is to leverage centralized Helpdesk and Monitoring service instances. For the **Helpdesk**, the service instance provided either by the EOSC EU Node (EEN) or by EGI through the EOSC Mesh project will enable the centralization of user support requests for the Data Terra Node. As the operational support teams are already in place, they will use the centralized ticketing system to manage user requests and provide effective support. Data Terra is also actively involved in the working group co-designing the EOSC Federation Helpdesk, with the following objectives: (i) Provide an overview of the Minimum Viable Product (MVP) for Federating Capability (FC-3); (ii) Define ticket management workflows; (iii) Define the escalation process; and (iv) Define the minimum information to be exchanged in cross-node tickets. For **Monitoring**, a similar approach will be adopted by relying on a dedicated instance provided through EOSC Mesh. Data Terra services, which already expose health-check endpoints, will be monitored from this centralized instance, providing a unified view of service availability and operational status across the node.

Significant progress has been achieved on the definition of the **Service Management System (SMS)** through the establishment of an internal working group dedicated to harmonising service management practices across the EOSC Node Data Terra. As part of this work, a common framework has been developed for the description of services and their associated Service Level Agreements (SLAs). To ensure consistency across the service portfolio, three standard SLA categories have been defined—Standard, High, and Critical—providing a common approach for specifying service commitments and operational expectations. In parallel, a common incident management framework has been established, including a harmonised severity classification. Severity 1 corresponds to a complete service outage affecting multiple users, a security breach, or data loss; Severity 2 refers to a significant service degradation impacting users while the service remains operational; and Severity 3 covers minor, non-critical issues with limited impact. These activities represent an important step towards implementing a consistent and EOSC-aligned Service Management System across the Node, supporting harmonised operations, user support, and future federation with EOSC service management processes.

The EOSC Node Data Terra **Landing Page** (www.earth-data.eu) is being developed as a single, coherent entry point providing integrated access to all services offered by the node through a user-centric interface. Its primary objective is to facilitate the discovery, access, and use of the Node’s resources while reducing fragmentation across the environmental research landscape. Beyond providing access to services, the landing page will showcase onboarded resources

through dedicated documentation and scientific use cases, demonstrating their value for researchers and other stakeholders. It will also act as a navigation hub, enabling cross-federation discovery and access to services, data sources, and research products in the domains of Earth system, biodiversity, and environmental sciences.

To date, a user needs and requirements analysis has been completed to capture stakeholder expectations and guide the platform design. The architecture of the landing page has been defined for its static content, while a test environment has been deployed for the catalogue frontend and successfully connected to the Resource Catalogue API. Development is currently focused on integrating dynamic EOSC functionalities. Additional activities include the implementation of multilingual support, the collection of stakeholder feedback to further improve the user experience, and the deployment of essential security measures covering backups, migration procedures, and production server configuration. The first operational release, planned for October 2026, will provide users with the ability to search and discover onboarded services, data sources, and research products through the Resource Catalogue (FC-2). The landing page will also provide access to the Helpdesk (FC-3), including contact mechanisms and user documentation, while service monitoring capabilities (FC-4) will be integrated through the EOSC monitoring framework. Scientific use cases will be linked to the catalogue to illustrate reproducible research workflows and demonstrate the added value of the available resources. User authentication will rely on the Data Terra Single Sign-On integrated with the EOSC AAI Federation (FC-1), enabling authenticated access to services directly from the landing page. More advanced functionalities, such as Research Product Accounting (FC-7) and Service and Resource Order Management (FC-8), are currently outside the scope of the October release and will be considered in subsequent development phases.

1.2 Horizon Europe projects contributing to the Node

EOSC-Connect
In EOSC-CONNECT, coordinated by CSC, Data Terra leads WP4, "EOSC Federating Capabilities Supporting Data Interoperability". Data Terra, through CNRS in coordination with IRD and INRAE as affiliated entities, brings complementary expertise in multidisciplinary technical interoperability and semantic artefacts and catalogues, respectively. The most significant contribution to Data Terra's operational development is around Semantic Interoperability with the Federation of Semantic Artefact Catalogues. This action will transform the OntoPortal federation prototype developed in FAIR-IMPACT project into a fully operational federated capability for EOSC by strengthening core technical services (Annotator, Recommender, and MOD-API), establishing a shared governance framework, and expanding the federation to additional research communities. It aims to the deployment of new capabilities on both AgroPortal and EarthPortal, directly enhancing the semantic capabilities of the Data Terra Node's artefact catalogue and improving its discoverability within the EOSC Federation. Data Terra will also benefit from WP4 developments on inter-node data sharing (repository-to-repository interoperability and integration with the EEN Resource Catalogue), as well as from AI-based discovery services, including ADEF (led by INRIA) and Open Knowledge Maps, which will increase the visibility and accessibility of the Node's resources for the European research community.
EOSC-Mesh
The project, coordinated by EGI, contributes to the EOSC Node Data Terra by reinforcing its role as the environmental node within the EOSC. Rather than developing new infrastructure, the project enhances Data Terra by integrating the EOSC Mesh federating capabilities, enabling seamless interoperability with other EOSC Nodes and facilitating cross-domain access to data, services, and computing resources. Data Terra also serves as one of the project's strategic adopters, validating these capabilities through concrete environmental science use cases. Data Terra also contributes scientific expertise through the "Environmental Impact of Volcanic Eruptions" use case, where heterogeneous Earth observation, in situ, and modelling data are integrated into reproducible workflows deployed on the Data Terra

platform and connected to EOSC services. Data Terra leads the "Ocean Monitoring for Risk Prevention and Biodiversity Conservation" use case, demonstrating the management and processing of large-scale marine observation data and validating EOSC Mesh capabilities for operational ocean science. Together, these contributions strengthen Data Terra as an interoperable EOSC Node, demonstrate the value of federated services for environmental research, and provide reusable workflows that support multidisciplinary collaboration across the EOSC ecosystem.

CDIF4EOSC

The contribution of the Data Terra node is structured around several complementary technical and operational axes. First, Data Terra serves as the Lead Beneficiary of Work Package 6 (WP6), which is responsible for the final systemic integration and deployment of the project's tools and services within the EOSC Federation. In this context, Data Terra acts as a pilot node for the operational demonstrators, showcasing the implementation of the CDIF4EOSC Playbook in real operational workflows, including the transformation of heterogeneous Earth observation and environmental datasets into FAIR Digital Objects (FDOs). Data Terra also contributes its recognized expertise in semantic interoperability and scientific variable description through the development of the I-ADOPT framework, while CNRS operates EarthPortal, an OntoPortal-based catalogue of semantic artefacts that provides a key infrastructure for metadata enrichment. In addition, Data Terra mobilizes its thematic data hubs, notably ODATIS for marine data and AERIS for atmospheric data, to validate CDIF4EOSC solutions in interdisciplinary use cases addressing oceanographic and climate research. Data Terra further contributes to policy and governance activities under WP6 by participating in stakeholder consultations with the EOSC Association to support the adoption of the CDIF4EOSC Playbook and the definition of sustainable governance and long-term operational models. Finally, Data Terra contributes to interoperability gap analysis by identifying bottlenecks in data ingestion and mapping interoperability requirements between EOSC nodes and European Data Spaces, including the Green Deal Data Space.

FLUID-AI

FLUID-AI, coordinated by CSIC, contributes significantly to the EOSC Node Data Terra by reinforcing its role as a flagship thematic node for Earth and environmental sciences within the EOSC Federation. Rather than creating new infrastructures, the project enhances the AI-readiness, interoperability and machine-actionability of the repositories and services operated by the Data Terra community through the development of a common Data and Models Liquidity framework. This includes the implementation of AI-ready repository guidelines, semantic enrichment of metadata, interoperable workflows, provenance and quality management mechanisms, and integration with EOSC Core services, enabling Data Terra resources to be more easily discovered, accessed and exploited by both researchers and AI applications. Data Terra also serves as one of the project's key validation environments, where representative Earth science use cases demonstrate the interoperability of data, models and AI services with other thematic EOSC Nodes and European Data Spaces. Data Terra brings the perspective of the Earth system community and leads activities related to interoperability with European Data Spaces and the implementation of Earth science use cases on the node. The EOSC Node Data Terra contributing organisations provide recognised expertise in semantic interoperability, ontologies, metadata harmonisation and AI-ready repositories. Finally, Ifremer strengthens the marine dimension of Data Terra by contributing its expertise in marine data infrastructures and repository interoperability, fostering seamless integration between marine and earth-related environmental data and reinforcing links with complementary EOSC thematic nodes such as DTO. Together, these contributions ensure that Data Terra evolves into a fully interoperable, trustworthy and AI-ready EOSC Node capable of supporting the next generation of Open Science and Digital Twin applications.

GenAI4Earth (Coordinated by CNRS / Data Terra)

Artificial Intelligence is transforming the way research is conducted, from literature review and hypothesis generation to data FAIRification and discovery, and multimodal data analysis. However, the adoption of AI in Earth System Science research raises pressing questions about

sovereignty, transparency, research integrity, trustworthiness and frugality. These concerns are central to the AI agenda of the EOSC Data Terra node, and in general the EOSC Federation, and to the future of national and European research infrastructures. Through the EOSC-GenAI4Earth project, and within the broader context of the AI4EOSC ecosystem, Data Terra and NFDI4Earth build a sovereign, generative and agentic AI environment to support the ESS research community. EOSC-GenAI4Earth will support the implementation of AI-driven FAIR data stewardship and federated research infrastructures tailored for AI in order to create a strong European foundation for next-generation digital research environments. The collaboration between Data Terra and NFDI4Earth focuses on implementing services to transform multi-modal Earth observation, in situ, and simulation datasets into AI-ready, machine-actionable research assets. The specific services will support AI-assisted FAIRification, AI-driven metadata generation, AI-enabled intelligent data discovery, and reproducible AI workflows to extract information from the data. The services are designed to operate natively within the EOSC Federation while remaining interoperable with national and European Research Infrastructures (ERICs). All services are built on the AI4EOSC framework. The EOSC-GenAI4Earth solutions will strengthen the EOSC Federation by providing interoperable, scalable and sustainable AI services that facilitate cross-domain scientific collaboration, accelerate data-driven discovery, and establish a common European ecosystem for trustworthy AI in ESS.

2. VALUE PROPOSITION

Data Terra brings to the EOSC Federation a unique combination of thematic expertise, operational research services, and interoperable digital infrastructure dedicated to Earth system, environmental and biodiversity sciences. As the French national research e-infrastructure in this domain, Data Terra federates five national data hubs and connects major European research infrastructures, enabling seamless access to multidisciplinary data, services, and scientific workflows. More broadly, Data Terra and its strategic collaboration with NFDI4Earth in the framework of the EOSC Node illustrate the added value that science-driven national research infrastructures bring to the EOSC Federation. Being closely connected to research laboratories and scientific communities, they are uniquely positioned to develop and continuously evolve services, data resources, and interoperability frameworks that directly address users' scientific requirements. This proximity enables rapid technical innovation, the validation of federating capabilities in operational research environments, and the deployment of reproducible cross-disciplinary workflows. The integration into the EOSC Federation will strengthen the European ecosystem by providing a trusted thematic node capable of supporting large-scale environmental research while ensuring interoperability with other EOSC Nodes.

A key value proposition of Data Terra Node lies in addressing several current challenges of the EOSC ecosystem. First, it reduces fragmentation by federating heterogeneous resources - including datasets, software, publications, services and semantic artefacts. Second, it contributes to the implementation and validation of EOSC Federating Capabilities, thereby supporting the operational deployment of the EOSC Federation. Third, Data Terra develops reproducible scientific workflows and cross-disciplinary use cases that integrate Earth observation, in situ measurements and numerical modelling, demonstrating the practical value of federated services for environmental research. Finally, through its expertise in FAIR data stewardship, semantic technologies, AI-ready repositories and metadata enrichment, the EOSC Node Data Terra contributes to improving the discoverability, interoperability and machine-actionability of research resources across EOSC.

The primary **beneficiaries** of these contributions are researchers, research infrastructures, universities and public research organisations working in Earth system, climate, biodiversity and environmental sciences. Public authorities and policymakers also benefit from improved access to high-quality environmental information supporting evidence-based decision-making in areas such as climate adaptation, biodiversity conservation, natural hazards and sustainable resource

management. More broadly, citizens and society benefit from strengthened Open Science practices, increased transparency, improved reproducibility of scientific results and enhanced support for research addressing major societal challenges and the Sustainable Development Goals.

As part of the EOSC Federation, Data Terra Node will better meet the needs of its user communities by providing a single-entry point to distributed environmental resources through a user-centric landing page, federated authentication, interoperable catalogues and harmonised support services. Researchers will be able to discover, access and combine resources from multiple EOSC Nodes through standard interfaces, while service providers will benefit from common operational services and increased visibility within the European research ecosystem.

The EOSC Node Data Terra operates at the interface between national infrastructures, European Research Infrastructures (ESFRI/ERICs), the Copernicus ecosystem, the AI European Factories, the emerging European Data Spaces and Destination Earth (DestinE), while providing access to advanced computing resources through the European HPC ecosystem. These connections ensure that EOSC services are interoperable with complementary European digital infrastructures while facilitating the uptake of artificial intelligence, digital twins, and data-intensive research across scientific domains.

Finally, as trusted national infrastructures coordinated within the EOSC Federation, Data Terra and NFDI4Earth contribute to strengthening European data sovereignty and the long-term preservation of scientific assets. Their distributed and coordinated approach reinforces Europe's capacity to preserve, curate, and provide sustainable access to Earth system data and services, while supporting international scientific collaboration and helping to build a resilient, interoperable, and trustworthy European infrastructure for environmental research.

3. USE CASES

01 Galaxy for Earth System Sciences (GESS)

The *Galaxy for Earth System Sciences* Use-Case aims to provide a robust and scalable computational framework to support earth system sciences within the EOSC Federation. This work will continue the one started within the EOSC project FAIR-EASE.

Use Case ID	Use Case Description	Federation Contributions & Value to Users
<i>01 Galaxy for Earth System Sciences (GESS)</i>	GESS Use Case aims to provide a robust, scalable, and interoperable computational framework to support Earth system sciences within the EOSC Federation. Building on the work initiated in the FAIR-EASE project, this UC extends collaboration with new partners (e.g., NFDI4Earth) and strengthens existing integrations (e.g., Pulsar). The goal is to democratize access to tools, workflows, and computational resources for environmental researchers while ensuring reproducibility, FAIRness, and cross-disciplinary collaboration. This use case focuses on:	Multi-node scientific workflows: We will connect the user environment across different EOSC nodes, leveraging high-performance computing (HPC) and cloud infrastructures to execute distributed workflows. Data Access & Management: We will connect the user environment with EOSC services, providing better access to data via APIs, S3 buckets, and direct links. Collaboration & Sharing: Facilitating interdisciplinary collaboration by improving Galaxy's interoperability and

	1. Enable researchers to process and analyse large-scale environmental datasets using Galaxy workflows. 2. Use, improve, and enhance the interoperability of the dedicated user environment <i>Galaxy for Earth System Sciences</i> - created within FAIR-EASE - to integrate EOSC resources. This is a joint work with Life Science Connect via ELIXIR both at a European and French level (integration of GESS with IFB the french ELIXIR node on Galaxy France). 3. Support new communities in bringing standardized tools, workflows, and training to the scientific community, ensuring accessibility and reproducibility of research. 4. Foster cross-domain analysis and collaboration within the Galaxy for Earth System Sciences. 5. Forests (Joint UC with NFDI4Earth): Integration of dedicated workflows for forestry data analysis (e.g., LiDAR data processing, 3D visualisations and specifications) in collaboration with NFDI4Earth. This includes: • Testing on diverse forest data (Data Terra's data from Theia hub) and improving the workflow • Sharing of best practices and reusable workflows via WorkflowHub. • Training and tutorials on the Galaxy Training Network. • Use of Pulsar to execute the workflow to test proof Pulsar nodes on HPC/cloud infrastructures (e.g., UCA, IFB, Data Terra's hub and partners).	connecting Galaxy users across different European research infrastructures. Training & User Support: Implementation of tutorials, webinars, and training sessions to assist researchers in effectively using Galaxy for Earth Sciences.
--	---	--

In- Scope

Core Functions

- A dedicated Galaxy user environment for Earth sciences.

- A support in developing & integrating New Tools & Workflows to meet the needs of researchers in environmental and Earth sciences.
- Galaxy integrated features such as AAI check-in, Data access & storage, Helpdesk, AI tools.
- A connected network of Distributed Computing Resources, integrating with services like One Data, and Pulsar.
- Interoperability with European (EOSC or other) and national services such as Blue-Cloud, and ENVRIHUB services.

Stakeholder Support

- Researchers in Earth System Sciences who require scalable computing resources.
- Research Infrastructures looking for integration of scientific workflows and connection with other IR such as NFDI4Earth (joint forests UC).

Integration

- EOSC Nodes & Federated Resources: The Use case will consist primarily to integrate resources from other EOSC services (e.g., compute, data storage, data access).
- Galaxy Interoperability: Ensure compatibility with other services and standards, such as WorkflowHub & RO-Crate, to allow sharing and reusability of best-practice workflows.
- Training & User Engagement: Collaborate with the Galaxy training community to provide ongoing user support.

Out of Scope:

Excluded Activities

- Direct funding of hardware or cloud infrastructure (relies on EOSC and external national and European funding sources).
- Development of new data storage solutions (leverages existing EOSC storage resources like One Data).
- Support for research domains outside of Earth system sciences.

Limitations

- Compute capacity depends on the availability of external collaborations (e.g., EGI, IFB, UCA, EuroHPC).
- Initial focus on European stakeholders, with potential expansion to a global audience in the long term.

Dependencies

- Successful enrolling of EOSC EU node services and standards
- Onboarding of EOSC nodes (Blue-Cloud, ENVRIHUB) and NFDI4Earth.
- Long-term funding sustainability, potentially requiring an ELIXIR-equivalent stamp for Earth sciences.

02 Volcano Space Observatory Use Case (VSO)

The Volcano Space Observatory use case showcases the integration of federated Earth observation services for near real-time volcanic hazard monitoring and analysis.

Use Case ID	Use Case Description	Federation Contributions & Value to Users
<i>02 Volcano Space Observatory (VSO)</i>	The <i>Volcano Space Observatory</i> Use-Case aims to evaluate new ways of handling data processing services in the Volcano Space Observatory (VSO). The VSO is a set of interactive web platforms dedicated to near real time, retrospective tracking and analysis of volcanic events, using Atmosphere and solid Earth data issued from ground-based, satellite measurements and	Scaling up the platform in the EOSC federation would help us solve technical bottlenecks while processing InSAR data. Data Access & Management: Evaluate easier and faster access to Copernicus datasets. Evaluate data processing on other European Federated infrastructures,

	models. Data coming from the marine domain is not supported at the moment but could be added based on the technical assessment carried out during this phase.	Evaluate access and processing in WekEO DIAS or OpenEO CDSE instances Multi-node scientific workflows: Access to marine data for integration in the platform through the DTOEOSC Node. The goal is to provide users with more accessible and faster services using interoperable APIs.
--	---	--

In-Scope:

Core Functions:

- Provide a single portal that merges all web services, documentation and materials developed over the historical projects.
- Testing dedicated data access on the Copernicus Data Space Ecosystem and/or CNES GEODES for COPERNICUS data (Sentinel), especially data extractions on areas of interest and remote processing,
- Testing automation of SO₂ Flux Calculation on OpenEO instance hosted by Copernicus Data Space Ecosystem.
- Evaluate access to ARGO data through the DTO EOSC Node,
- Extend data supported in the platform with ACTRIS ERIC datasets (Historical and NRT) to cover the European geographical area: ground-based aerosol Lidar backscatter profiles, ground-based chemical composition of aerosols (ACSM), Evaluate the potential contribution to the Copernicus Atmosphere Monitoring Service (CAMS) by adding satellite-derived gas emission rate data for worldwide volcanoes from the Volcano Space Observatory to the Global Volcanic Emissions dataset (CAMS-GLOB-VOLC). Currently, the CAMS-GLOB-VOLC dataset only gathers ground-based SO₂ measurements from the global Network for Observation of Volcanic and Atmospheric Change (NOVAC).

Stakeholder Support:

- Researchers in Earth System Science (Atmosphere and Solid Earth academics, Climate scientists).
- Teachers, students.
- Scientific and technical staff working in Volcano Observatories.
- Institutions in charge of air quality monitoring.
- Institutions in charge of air traffic safety (eg. VAAC: Volcanic Ash Advisory Centers).
- World Climate Research Programme via the APARC (Atmospheric Processes And their Role in Climate) core project that coordinates the evaluation of the impact of volcanism on climate.

Integration:

- Documentation of the existing services.
- Development of training material and tutorials.
- Development of a standardized REST API that would allow triggering computation processes, into a machine-to-machine operable way.

Out of Scope

Excluded Activities:

- Operational InSAR processing on new European infrastructures. Only the evaluation of the feasibility and potentially the prototyping of the method is conducted during this phase.
- By its cross-domain nature, the VSO platform is a matter of interest for the EPOS ERIC via the Volcano Observations Thematic Core Service. Links with this research infrastructure are currently being constructed but still need to be consolidated.

Limitations:

- Access to Sentinel data on the Copernicus Data Space Ecosystem is currently limited in performance and volume. Scaling up of the VSO would necessitate the limitations are lifted or reduced.
- Running containerized processes on the WeKEO DIAS or OpenEO CDSE instance. A study has already been conducted (2025-2026) by porting an algorithm on this platform and some limitations were highlighted. In particular, WekEO computing resources are not free by default. Either a framework with WekEO should be found to use the platform at no cost, or funding solutions should be provided.

Dependencies:

- Successful on-boarding and enrolling with DTO EOSC node.
- Internal human and computing resources availability during the development phase, given that no dedicated funding is provided.
- Long-term funding sustainability.

03 Cross-Node workflows for qualification, calibration and validation of Biogeochemical ARGO float observations - Second wave Nodes: Turkey and EGI

This use case provides a reproducible and scalable Galaxy environment for the qualification, calibration and validation (QCV) of Biogeochemical (BGC) Argo float observations, enabling standardized workflows and FAIR-compliant validated datasets. The same workflows support large-scale reprocessing and the production of value-added products for initiatives such as Copernicus Marine and the Digital Twin of the Ocean.

Use Case ID	Use Case Description	Federation Contributions & Value to Users
<i>03 Biogeochemical ARGO float observations</i>	Biogeochemical (BGC) Argo floats provide essential observations for monitoring ocean biogeochemistry, climate change, ocean health and marine ecosystem dynamics. Ensuring the quality, consistency and traceability of these observations before their integration into international repositories and downstream services remains a major challenge for the scientific community. The Data Terra Node, through its marine data and services ecosystem and the Galaxy Earth System platform, aims to provide a reproducible and scalable environment enabling BGC Argo Principal Investigators (PIs) to qualify, calibrate and validate their observations using harmonised workflows. The service supports the ingestion of float profiles, application of quality-control procedures, calibration methods, visual inspection and generation of FAIR-ready validated datasets. Beyond supporting individual PIs, the same workflows can be reused for large-scale generation of derived products, reprocessing activities, Copernicus	Reproducible qualification, calibration and validation workflows for BGC Argo observations. Shared execution environment accessible to distributed scientific communities. Federated access to cloud and HPC resources through EOSC. Improved traceability, consistency and FAIRness of validation procedures. Capacity to scale from individual PI validation activities to basin-scale and global reprocessing campaigns. Reusable workflows supporting Copernicus Marine and Digital Twin of the Ocean initiatives.

	Marine services and future Digital Twin of the Ocean applications. The use case also demonstrates the value of the EOSC Federation by combining thematic scientific services operated by Data Terra with federated compute and storage resources provided by the EGI Node and HPC resources contributed through the Turkish Node. This enables scientists to access advanced processing capabilities without requiring local infrastructure while preserving reproducibility, transparency and FAIR principles.	
--	---	--

In- Scope

Core Functions

- Galaxy Earth System platform operated within the Data Terra ecosystem.
- Qualification, calibration and validation workflows for BGC Argo observations.
- FAIR workflow publication and sharing.
- Integration with EOSC Federation services.
- Cross-node execution using cloud and HPC resources.
- Generation of validated datasets and derived nitrate products.

Stakeholder Support

- BGC Argo Principal Investigators.
- Argo data assembly and validation centres.
- Marine biogeochemistry researchers.
- Copernicus Marine and Digital Twin of the Ocean communities.

Integration

- Data Terra: scientific workflows, domain expertise and service operation.
- EGI Node: federated cloud, storage and EOSC integration services.
- Turkish Node: HPC resources exposed through the federation.
- EOSC AAI and federated service catalogue integration.

Out of Scope:

Excluded Activities

- Development of new sensor calibration algorithms.
- Development of third-party scientific libraries.
- Operational management of the global Argo network.

Limitations

- Initial focus on BGC Argo qualification, calibration and validation workflows.
- Operational service deployment beyond the EOSC Federation build-up phase is outside the current scope.

Dependencies

- Availability of BGC Argo observations and reference datasets.
- Availability of cloud resources from EGI.
- Availability of HPC resources from the Turkish Node.
- Continued interoperability between Galaxy and EOSC Federation services.

04 GRAVITY UC1-2:

The first **use case** focuses on **Holocene fossil pollen records** (c. 11,700 years BP to the present), which document past vegetation, biodiversity, and climate variability. Aggregated and selected by Tallinn University of Technology (TalTech), the dataset provides better-than-centennial-resolution records from central and northern Europe. These well-dated, quality-controlled

samples document plant community composition and dispersal through time. Because the fossil taxa can be linked to Nearest Living Relatives (NLRs), they can be compared directly with modern plant distributions and environmental datasets. The records enable the reconstruction of vegetation dynamics and the derivation of quantitative climate parameters (e.g. temperature and precipitation). As part of the EU Horizon Europe *Past2Future* project, which aims to improve climate model performance, this dataset illustrates how high-resolution palaeoecological data can complement modern observations in analysing long-term environmental change.

A second use case, the **geological use case**, will focus on stratigraphically constrained mineral, rock, or fossil material relevant to solid Earth applications. Potential examples include specimens documenting past hazard events or geological processes, such as those described in the CETAF ESG publication on hazards and disasters in the geological record (Tilley et al., 2019). These materials demonstrate how collection-based geological evidence can inform hazard assessment, environmental change, and resource analyses. The final dataset will be defined in consultation with Data Terra partners to ensure alignment with solid Earth priorities.

<i>Use Case ID</i>	<i>Use Case Description</i>	<i>Federation Contributions & Value to Users</i>
<i>UC1</i>	Integration of quality-controlled Holocene and Late Glacial fossil pollen records from Estonia, Sweden and the Baltic region through GeoCAsE and the EOSC Data Terra Node. The use case demonstrates how palaeoecological specimen and observation data can be harmonised using FAIR metadata standards and linked with modern Earth Observation datasets to reconstruct long-term vegetation dynamics, biodiversity changes and climate variability.	Demonstrates interoperability between natural science collections and Earth Observation platforms. Provides researchers with access to harmonised palaeoenvironmental datasets spanning thousands of years, extending the temporal depth of environmental analyses and supporting climate modelling, ecosystem reconstruction and biodiversity research. The workflow establishes a reusable model for integrating palaeoecological collections within the EOSC Federation.
<i>UC2</i>	Integration of high-resolution Holocene tree-ring proxy data from northern Fennoscandia, including ring width, stable carbon and oxygen isotope ratios, blue-light intensity and tree-ring density. These annually resolved records reconstruct past climate variability over the last 7,500 years and preserve signatures of major volcanic eruptions through abrupt reductions in tree growth and changes in stable isotope composition. The use case demonstrates how dendroclimatological records can be linked with Greenland ice-core archives, North Atlantic climate records and Earth Observation services through GeoCAsE and the EOSC Data Terra Node.	Demonstrates interoperability between specimen-based collections, palaeoclimate databases and EOSC services. Researchers gain access to harmonised, annually resolved climate proxy data for reconstructing past temperature and hydroclimate variability, identifying the environmental impacts of volcanic eruptions, validating Earth system and climate models, and investigating abrupt climate change. The workflow provides a scalable framework for integrating high-resolution climate proxy data within the EOSC Federation.

4. External Dependencies & Key Risks

External Dependencies & Risks	Actions	Deadline
Helpdesk and Monitoring FC are still under development, potentially delaying their integration into the Data Terra Node.	Continue active participation in EOSC working groups and Horizon Europe projects (EOSC Mesh); integrate services as soon as stable production instances become available; maintain local operational capabilities until federation services are operational.	October 2026 (initial integration); continuous alignment thereafter.
Evolution of EOSC Federation specifications may require adaptations of technical implementations.	Participate in EOSC governance and technical working groups; maintain modular architecture allowing rapid adaptation to updated specifications and Federating Capabilities.	Continuous throughout the project.
Interoperability challenges arising from heterogeneous metadata models and services by multiple providers.	Implement unified DCAT-based catalogue model, semantic interoperability services, automated metadata validation, and quality control before exposing metadata to the EOSC EU Node.	Resource Catalogue operational by October 2026; continuous improvements afterwards.
Distributed service operation across multiple national providers may affect consistency of support and operations.	Harmonise operational procedures, centralise user support through the EOSC Helpdesk when available, and deploy common monitoring and service management practices.	Initial implementation by October 2026; operational consolidation during 2027.
Availability of technical resources and long-term sustainability beyond project funding.	Build developments on long-term national infrastructures and institutional commitments (Gaia Data – FR); align with national Research Infrastructure Roadmap.	Continuous.
Compliance with EOSC operational, security and data protection requirements (GDPR, AAI, incident response, metadata quality).	Maintain compliance with the AARC Blueprint Architecture, GDPR requirements, EOSC operational policies and security procedures; perform regular validation of metadata, authentication services, backups, and production deployments.	Before production release (October 2026) and continuously thereafter.
Cybersecurity and operational risks affecting service availability or integrity.	Implement secure production environments, regular backups, monitoring, incident response procedures, controlled access	Before production release and continuous operations.

	management, and vulnerability management.	
Dependencies on external strategic initiatives (AI Factories, European Data Spaces, Destination Earth, HPC Federation) whose roadmaps may evolve independently.	Maintain active collaboration with partner initiatives and ensure interoperability through open standards and modular interfaces rather than project-specific integrations.	Continuous.

5. Contributions [Deliverables (including documentation)]

ID	Deliverable Name	Deliverable Description	Deliverable Owner	Target date
D1	Federating Capabilities implementation plan	Technical and operational roadmap describing the implementation, deployment, and adoption of the EOSC Federating Capabilities, including dependencies, milestones, responsibilities, and integration with the EOSC EU Node.	Data Terra	September 2026
D2	Resources catalogue	Operational deployment of the EOSC Data Terra Resource Catalogue.	Data Terra	October 2026
D3	Landing Page	Deployment of the EOSC Node Data Terra landing page providing a single-entry point to node services, resource discovery, documentation, user support, and access to federated EOSC capabilities.	Data Terra / NFDI4Earth	October 2026
D4	Data Terra - NFDI4Earth MoU signature	Signature of the Memorandum of Understanding establishing the strategic collaboration framework between Data Terra and NFDI4Earth for the development of a transnational thematic EOSC Node.	Data Terra / NFDI4Earth	October 2026
D5	GRAVITY UCs demonstrators	Deployment and validation of the GRAVITY scientific use case demonstrators.	Data Terra / GRAVITY partners	November 2026
D6	UCs implementation plan	Definition and continuous update of the implementation roadmap for scientific use cases, including priorities, technical requirements, dependencies, milestones, and validation activities across participating partners.	Data Terra	Ongoing

D7	User engagement plan	Implementation of user engagement activities, including stakeholder consultation, user onboarding, training, dissemination, documentation, and collection of feedback to support continuous improvement of the EOSC Data Terra Node.	Data Terra / NFDI4Earth	Q1 2027
----	----------------------	--	-------------------------	---------

6. Timing and Milestones

ID	Milestone Description	Target Delivery Date
M1	Federating Capabilities Implementation Plan Approved	September 2026
M2	Data Terra Resource Catalogue Operational	October 2026
M3	EOSC Data Terra Landing Page Online	October 2026
M4	Data Terra–NFDI4Earth Partnership Established (Node governance)	October 2026
M5	First EOSC Federating Capabilities Operational	October 2026
M6	GRAVITY Demonstrators Validated	November 2026
M7	User Feedback Assessment Completed for the next Node release	March 2027

7. Contacts

Role	Name	Email
Coordinator	Alessandro Rizzo	alessandro.rizzo@ird.fr
Deputy Coordinator	Christelle Pierkot	christelle.pierkot@cnrs.fr
Operation Manager	Erwan Bodéré	erwan.bodere@ifremer.fr
Deputy Operation Manager	Marie Jossé	marie.josse@ifremer.fr
Security Officer	Sabine Avellaneda	sabine.avellaneda@data-terra.org
Communication Officer	Marina Ripon	marina.ripon@ipsl.fr
Legal Officer	TBD	