Ref. Ares(2024)4684399 - 28/06/2024

ANERIS

Operational Sensing Life Technologies for Marine Ecosystems

`

Deliverable D5.3 – EOSC Integration Plan

Lead Beneficiary: EGI Foundation

Author/s: Sebastián Luna-Valero, Rubén Díez, Ivan Rodero, Rune Lagaisse, Gergely Sipos

28th June 2024

Funded by
the European Union

Views and opinions expressed are those of the author(s) only and do not necessarily reflect those of the European Union or the European Commission. Neither the EU nor the EC can be held responsible for them.

Prepared under contract from the European Commission

Grant agreement No. 101094924

EU Horizon Europe Research and Innovation action

Project acronym:	**ANERIS**
Project full title:	**operAtional seNsing lifE technologies for maRIne ecosystemS**
Start of the project:	January 2023
Duration:	48 months
Project coordinator:	Jaume Piera
Deliverable title:	EOSC Integration Plan
Deliverable n°:	D5.3
Nature of the deliverable:	Report
Dissemination level:	Public
WP responsible:	WP5
Lead beneficiary:	EGI Foundation
Citation:	Luna-Valero, S., Díez, R., Rodero, I., Lagaisse, R., Sipos, G. (2024). *EOSC Integration Plan*. Deliverable D5.3 EU Horizon Europe ANERIS Project, Grant agreement No. 101094924
Due date of deliverable:	M18
Actual submission date:	M18

Deliverable status:

Version	Status	Date	Author(s)
0.1	Draft	30th May 2024	Sebastian Luna-Valero, Gergely Sipos EGI Foundation
0.2	Draft	10th June 2024	Sebastian Luna-Valero (EGI), Rune Lagaisse (VLIZ), Gergely Sipos (EGI)
1.0	Final	21st June 2024	Ivan Rodero (Quanta), Sebastian Luna-Valero (EGI)

The content of this deliverable does not necessarily reflect the official opinions of the European Commission or other institutions of the European Union.

Table of Contents

Executive summary **4**
List of Abbreviations **5**
1. EOSC and its offering **6**
1.1 What is EOSC? 6
1.2 How is it being implemented? 6
1.3 Researchers: How to participate? 7
1.4 Contributors: How to participate? 10
1.5 Transition to a new landscape 11
2. ANERIS and its offering **12**
2.1 Key Exploitable Results (KERs) 12
2.1.1 KER 1: ANERIS Technologies 12
2.1.2 KER 2: ANERIS Training 13
2.1.3 KER 3: ANERIS Co-Design 13
2.1.4 KER 4: ANERIS Cyberinfrastructure 13
2.1.5 KER 5: ANERIS Community 13
2.2 Operational Marine Biology data products (OMB) 14
2.2.1 Species diversity map 14
2.2.2 Intraspecific genetic variations map 15
2.2.3 Non-indigenous species map 16
2.3 Shortlisted ANERIS artifacts 17
3. EOSC integration plan **19**
3.1 Software 19
3.2 Documents (reports, presentations, training material) 20
3.3 Methodologies 20
3.4 Publications 21
3.5 Data Sources 22
3.6 Services 23
3.6.1 EOSC AAI 23
3.7 EOSC Nodes 24
Conclusions **24**
References **26**

Executive summary

The ambition of the European Open Science Cloud (EOSC) is to provide European researchers, innovators, companies and citizens with a federated and open multi-disciplinary environment where they can publish, find and reuse data, tools and services for research, innovation and educational purposes. This deliverable presents the plan for ANERIS to make available its results in EOSC. EOSC is currently in radical reimplementation, with several of its core concepts being redesigned. Following a phase of implementation through EC projects between 2016-2023, EOSC in 2024 is transitioning from a rather monolithic structure to a 'system of systems' architecture that integrates both EC, national and thematic funding and resources. The elements of the new EOSC are called Nodes, and the EC through a procurement activity is assembling the first Node, the EOSC EU Node, since February 2024. Considering this landscape this document consists of 3 sections:

Section 1 presents EOSC and how it is being implemented in parallel with ANERIS as a partnership of 'Nodes'. The section outlines the options that developers/contributors have to engage with EOSC either through the EU Node, or through national and thematic Nodes that are expected to be established in the future.

Section 2 summarizes those results from ANERIS that fit into the scope of EOSC and are therefore subject to integration. These services include 5 Key Exploitable Results that consist of both technological and human support achievements, and 3 Operational Marine Biology data products that are biodiversity information systems for systematic and long-term routine measurements of the ocean and coastal life and for their rapid interpretation and dissemination.

Section 3 provides suggestions on how ANERIS could make its results available in EOSC: The Zenodo repository (part of the EU Node) is recommended for Software and Document/Slides outputs; the 'Open Research Europe' system is recommended for scientific publications; the Resource Hub section of the EU Node is recommended for scientific data and online services. Also, depending on the future landscape of thematic nodes ANERIS should consider sharing its output via one/more thematic Nodes (e.g. a Marine Sciences Node) instead of the EU Node.

Because EOSC is in rapid evolution this section will have to be revisited and further revised as ANERIS and EOSC move forward. There are two scheduled submissions for this deliverable: one in M18 of the project (June 2024), and another one in M36 (December 2025). This version is for M18 and therefore is an initial integration plan. An updated version will be prepared in 18 months from now.

List of Abbreviations

AAI – Authentication and Authorization Infrastructure

AIES-MAC – Automatic Information Extraction System for MACro-organisms

AIES-PHY – Automatic Information Extraction System for PHYtoplankton images

AIES-ZOO – Automatic Information Extraction System for ZOOplankton images

AMAMER – Advanced multiplatform App for Marine lifE Reporting

AMOVALIH – Advanced Marine Observations VALidation-Identification system based on Hybrid intelligence

ARMS – Autonomous Reef Monitoring Structures

ATIRES – Automatic underwaTer Image REstoration System

AWIMAR – Adaptive Web Interfaces for MARine life reporting, sharing and consulting

DIH – Digital Innovation Hub

DOI – Digital Object Identifier

EMUAS – Expandable Multi-imaging Underwater Acquisition System

ENVRI – Environmental Research Infrastructures

EOSC – European Open Science Cloud

FAIR – Findable, Accessible, Interoperable, and Reusable

KER – Key Exploitable Result

MARGENODAT – workflows for the MARine GENOmics DAta managemenT

MBON – Marine Biodiversity Observation Network

NANOMICS – NAnopore sequeNcing for Operational Marine genomICS

OMB – Operational Marine Biology

RI – Research Infrastructure

RoP – Rules of Participation

VRE - Virtual Research Environment

1. EOSC and its offering

This first section provides a high level overview of the European Open Science Cloud (EOSC), what is the implementation plan, how can researchers benefit from EOSC and how can contributors take part in the building of EOSC.

1.1 What is EOSC?

The ambition of the European Open Science Cloud, known as EOSC[1], is to develop a **Web of FAIR Data and Services** for science in Europe.

The European Open Science Cloud will enable researchers to find, create, share and reuse all forms of digital knowledge – such as publications, data and software – leading to new insights and innovations, higher research productivity and improved reproducibility in science

The vision for EOSC is to put in place a system in Europe to find and access data and services for research and innovation. This is to help researchers store, share, process, analyse and reuse FAIR research outputs within and across disciplines and borders.

1.2 How is it being implemented?

EOSC builds on existing infrastructure and services supported by the European Commission, Member States and research communities. It brings these together in a federated **system of systems** approach, adding value by aggregating content and enabling services to be used together.

This environment will operate under well-defined conditions to ensure trust and safeguard the public interest. Expectations of service providers and users will be made explicit to ensure appropriate behaviour.

Connecting data repositories and services will be instrumental for Open Science to progress in Europe. For this, the **EOSC Federation**[2] of nodes is being created (see Figure 1.2.1).

[1] European Open Science Cloud: https://eosc.eu/eosc-about/
[2] EOSC Federation: https://eosc.eu/building-the-eosc-federation/

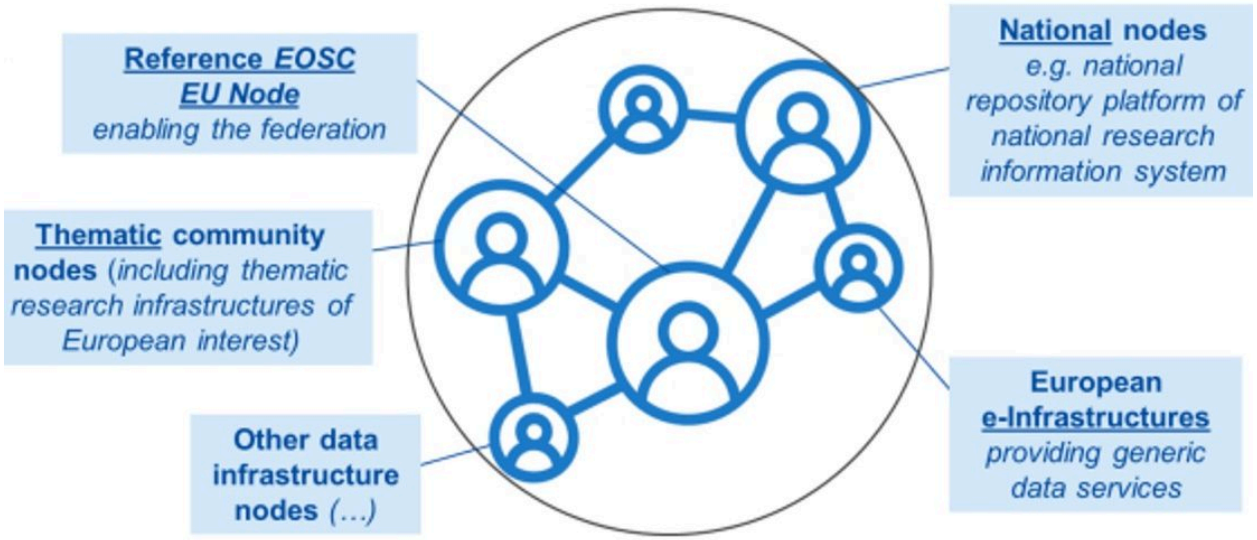

Figure 1.2.1. *The EOSC Federation connects multiple EOSC Nodes together*

The EOSC Federation will consist of multiple **EOSC Nodes** that are interconnected and can collaborate to share and manage scientific data, knowledge, and resources within and across thematic and geographical research communities. The EOSC Nodes will be entry points for users to the EOSC Federation, with each node offering its own and possibly third-party services, including data reposing and accessing services.

The **EOSC EU Node**[3] is the first of the EOSC Federation (see Figure 1.2.1). However, 1) the minimum requirements of an EOSC Node and 2) the enrollment of eligible Nodes in the EOSC Federation are still being discussed at the time of this writing[4].

1.3 Researchers: How to participate?

At the time of this writing, the best entry point for researchers that want to benefit from EOSC is to use the EOSC EU Node. The EOSC EU Node is a gateway to a world of collaborative research, no matter your field. It provides a one-stop shop for researchers across all disciplines, offering powerful tools and resources to streamline data-driven endeavours. With EOSC EU Node, it is possible to::

- Find and use existing data: Easily discover and access data sets relevant to your research even from different disciplines and locations.
- Store and share data: Securely store research data and effortlessly share it with colleagues across the globe.
- Analyse data: Access powerful computing tools and software to analyse data and make groundbreaking discoveries.

[3] The EOSC EU Node: https://open-science-cloud.ec.europa.eu/
[4] Towards a fully fledged EOSC Federation

- Publish findings: Share research results with ease, following the principles of Open Science.

The EOSC EU Node offers access to a set of managed services[5] which has been procured by the European Commission and are made available via the front office (Figure 1.3.1). In short, you will find application services for data management:

- A **File Sync and Share** facility to maintain consistent data sharing across your multiple work environments.
- A **Notebook** environment to create and share documents with real-time code execution.
- A **Large File Transfer** function to reliably send files to colleagues.

Information Technology (IT) oriented users will also find infrastructure services (Figure 1.3.1):

- Access to **Virtual Machines** to support processing, analytics, and other data and compute-intensive use cases.
- A **Cloud Container Platform** to streamline the execution of cloud-native research workflows.
- A **Bulk Data Transfer** system to efficiently move large datasets.

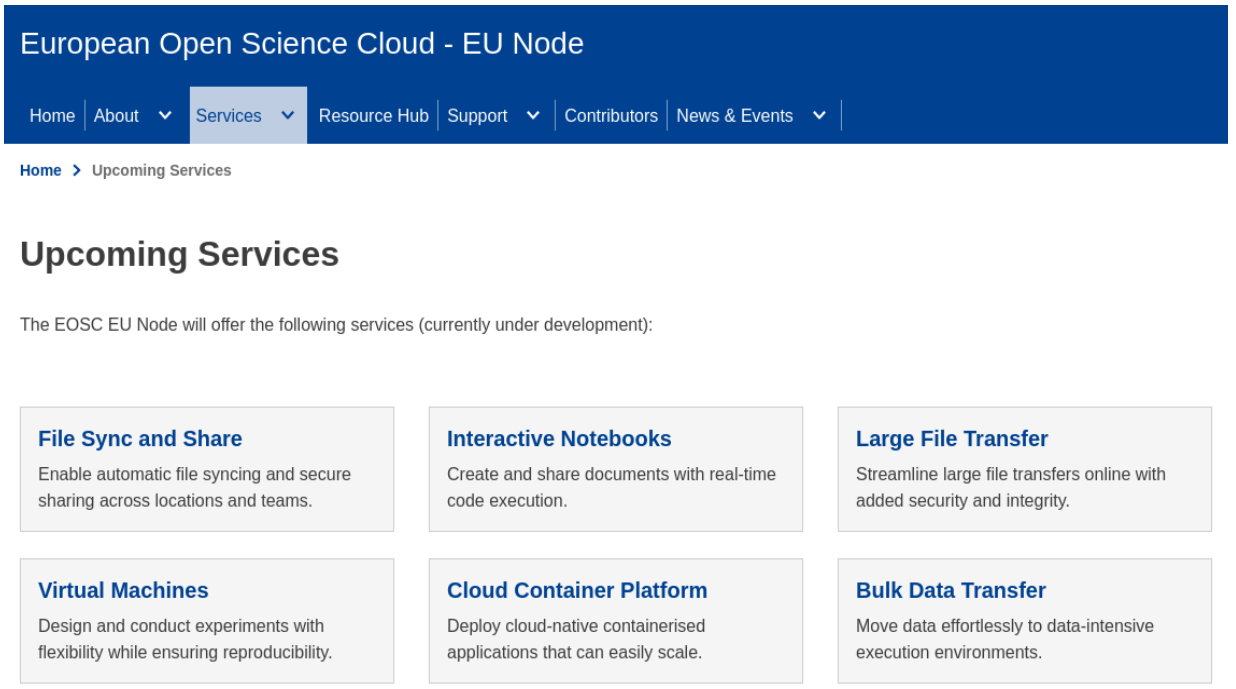

Figure 1.3.1. *Set of managed services procured by the European Commission.*

[5] Services offered by the EOSC EU Node: https://open-science-cloud.ec.europa.eu/services

These services are provided for free by the European Commission. Detailed Access Policy is subject to apply. Information about usage and accounting will be collected and transparently shown to users.

Moreover, the platform offers access to a **Resource Hub**[6] (Figure 1.3.2) which includes metadata collected from the broader EOSC ecosystem such as:

- Research data, publications, software published in EOSC and beyond
- Services offered by the broader research community
- Support and learning material to help you practice FAIR and Open Science
- Deployment tools and recipes that enable execution of data science workflows (coming in Q3 2024)

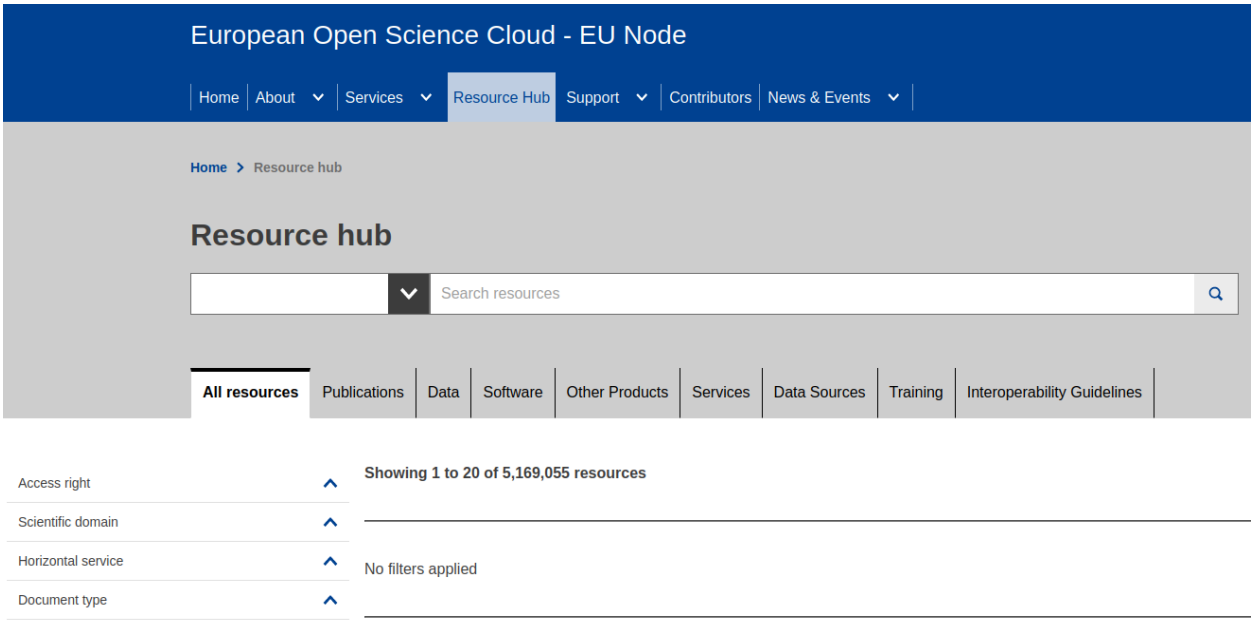

Figure 1.3.2. *Resource Hub offered by the EOSC EU Node*

Relevant services for ANERIS like the iMagine AI platform for aquatic sciences[7] (Figure 1.3.3) and the Global Record of Stocks and Fisheries[8] can be found using the Resource Hub at the EOSC EU Node.

[6] Resource Hub offered by the EOSC EU Node: https://open-science-cloud.ec.europa.eu/resources/
[7] iMagine AI platform for aquatic sciences
[8] Global Record of Stocks and Fisheries

Figure 1.3.3. *iMagine AI platform for aquatic sciences entry in the Resource Hub.*

The EOSC EU Node is expected to be fully operational by the end of the third quarter of 2024 (Q3 2024), with a phased introduction of its services. This will be accompanied by the introduction of additional features and services in the future.

For more information about the EOSC EU Node, please visit its FAQ[9] section on the website.

1.4 Contributors: How to participate?

Although the Rules of Participation (RoP) are still to be defined, those who want to help build EOSC can do so in these ways:

1. Contribute with curated and FAIR research outputs to the Resource Hub of the EOSC EU Node ensuring discovery and findability (Figure 1.3.2)
2. Onboard selected research outputs directly to the EOSC EU Node as an integrated infrastructure provider hooking up to the EOSC EU Node core capabilities
3. Establish an autonomous EOSC Node, considering the reference implementation of the EOSC EU Node, and become part of the EOSC federation.

So far this is the list of contributors categories identified by the EOSC EU Node:

- **Repositories**: Offer curated and FAIR research products with unique identifiers and minimum semantic interoperability (such as publications, data, software, etc.) that are as open as possible and as closed as necessary.
- **Research Infrastructures**: Contribute specialised knowledge, tools, applications, and datasets, particularly in niche or cutting-edge areas of science.

[9] EOSC EU Node FAQ: https://open-science-cloud.ec.europa.eu/support/frequently-asked-questions

- **Technology and Infrastructure Providers**: Offer infrastructure and/or platform services following the cloud-based delivery model, that can form the European backbone of computational and data storage capabilities.
- **Software Developers and IT Professionals**: Develop the models, tools, and applications that facilitate data management, analysis, collaboration, and other essential research activities within the EOSC ecosystem.
- **Industry Partners**: Engage with the EOSC EU Node for collaborative research activities spanning across public and private sectors and for potentially bridging the gap between scientific research and industrialised innovation.

The EOSC EU Node will offer in the future the possibility for Contributors to utilise some core capabilities “as a Service”. Those capabilities will be:

- **EOSC Authentication and Authorisation Infrastructure (AAI)**: A federated trust and identity management framework for EOSC.
- **Resource Catalogues and Registry Services**: Seamless connection and discovery of research objects and catalogues.
- **Application Workflow Management**: A workflow tool to compose and orchestrate infrastructure resources federated into EOSC.
- **Monitoring and Accounting**: Transparent monitoring and accounting information across EOSC.
- **Helpdesk**: Part of the Service Management System proposed for EOSC as a common framework for operationalised environments.

More details about this outsourcing model will soon be available. For more information on how to contribute, see the dedicated section[10] on the website.

Contributing to the EOSC EU Node presents a unique opportunity for you to be part of shaping the future, i.e., to have a say in the development of policies, standards, and best practices that shape the emerging EOSC federation. By sharing and integrating your resources in EOSC EU Node, not only do you increase their visibility within the European and global research communities and foster cross-disciplinary collaborations that were not possible before, but you can also make a stronger case for a more stable base of support and justify future investments.

1.5 Transition to a new landscape

The EOSC EU Node (commissioned in April 2024[11]) is the predecessor of the previous EOSC Portal[12] and EOSC Marketplace[13] (decommissioned in March 2024). This deliverable presents an integration plan with EOSC considering the new landscape with the EOSC EU Node.

[10] Contributors to the EOSC EU Node: https://open-science-cloud.ec.europa.eu/contributors
[11] Commission Announces the EOSC EU Node’s Web Presence
[12] Decommissioned EOSC Portal: https://eosc-portal.eu/
[13] Decommissioned EOSC Marketplace: https://marketplace.eosc-portal.eu/

2. ANERIS and its offering

This Section presents the artifacts produced by the ANERIS project. In particular Section 2.1 describes the Key Exploitable Results identified in (Salazar Forn, X., 2023). Section 2.2 introduces the Operational Marine Biology (OMB) data products from ANERIS, and Section 3.3 shortlists the outputs from ANERIS that are better suited to become part of EOSC.

2.1 Key Exploitable Results (KERs)

A **Key Exploitable Result (KER)** is an identified main interesting result of the project which has been selected and prioritised due to its high potential to be "exploited" – meaning that others will make use of it and derive benefits. See Section 2.1 in (Salazar Forn, X., 2023) for more details.

Each of the KERs identified in (Salazar Forn, X., 2023), Section 3.1, is described in a dedicated subsection below.

2.1.1 KER 1: ANERIS Technologies

The technologies developed in ANERIS are classified in 3 categories: Genomics, Imaging and Citizen Science.

There are 3 technologies specific for Genomics: NANOMICS, MARGENODAT and SLIM-2.0.

- NANOMICS: Nanopore sequencing for operational marine genomics.
- MARGENODAT: workflows for the marine genomics data management.
- SLIM-2.0: A VRE for genomic data analysis.

There are 5 technologies related to Imaging:

- AIES-ZOO: Automatic information extraction system for zooplankton images.
- AIES-PHY: Automatic information extraction system for phytoplankton images.
- AIES-MAC: Automatic information extraction system for macroorganisms.
- ATIRES: Automatic underwater image restoration system.
- EMUAS: Expandable multi-imaging underwater acquisition system.

There are 3 technologies for Citizen Science:

- AMAMER: Advanced multi-platform app for marine life reporting.
- AWIMAR: Adaptive web interface for marine life reporting, sharing and consulting.
- AMOVALIH: Advanced marine observations validation-identification system based on hybrid intelligence.

For more information about ANERIS technologies, please visit the ANERIS website[14].

[14] Technologies in the ANERIS website: https://aneris.eu/technologies

Out of the total 11 technologies described above, all of them are software products except for NANOMICS and EMUAS. On the one hand, NANOMICS is mostly about finding, developing, and optimizing laboratory protocols for Nanopore sequencing, so the end product is a document with recommendations for protocols and best practices. On the other hand, EMUAS is a multi camera system built with commercial off-the-shelf components and can be placed within a few hundred meters of shore (or on an existing cabled observatory). It is a cost-effective solution that can be used independently, or integrated into existing Research Infrastructures (RIs), expanding their capabilities.

2.1.2 KER 2: ANERIS Training

This KER gathers all the training material produced by the ANERIS project. This can take the form of a user guide, presentation, video recording, etc.

For example consider the workshop organized by the ANERIS project about “AI Basics for Image Processing” (for more details, see the ANERIS website[15]). The training material produced in this workshop, along with those from future workshops, are categorized under KER 2 and will be collected and shared with the general public to maximise its reuse.

2.1.3 KER 3: ANERIS Co-Design

ANERIS is employing a co-design methodology that helps partners in the project with the development and validation of technological solutions. This methodology is brought to the ANERIS project thanks to our consortium partner Science for Change[16].

Science for Change is working on a document to capture the co-design methodology in ANERIS and will publish it later in the year. This publication will be captured and presented in the next iteration of this deliverable.

2.1.4 KER 4: ANERIS Cyberinfrastructure

The ANERIS cyberinfrastructure was described in (Luna-Valero, S., 2023) and later validated in (Díez, R., 2023). This cyberinfrastructure consists of computing and storage resources distributed around Europe (cloud, fog and edge layers) with the aim to enable ANERIS services and workloads that will streamline the creation of Operational Marine Biology data products.

2.1.5 KER 5: ANERIS Community

The ANERIS community consists of researchers, RI operators, start-ups, innovative ecosystems and marine related DIHs, citizen scientists, etc. Activities from the community helps with the validation of the ANERIS cyberinfrastructure and its exploitation for the successful outcome of the project.

[15] ANERIS Workshop: https://aneris.eu/news/looking-back-aneris-workshops
[16] Science for Change: https://scienceforchange.eu/

2.2 Operational Marine Biology data products (OMB)

ANERIS proposes to implement the concept of **Operational Marine Biology (OMB) data products**, understood as a biodiversity information system for systematic and long-term routine measurements of the ocean and coastal life, and their rapid interpretation and dissemination. (see Figure 2.2.1).

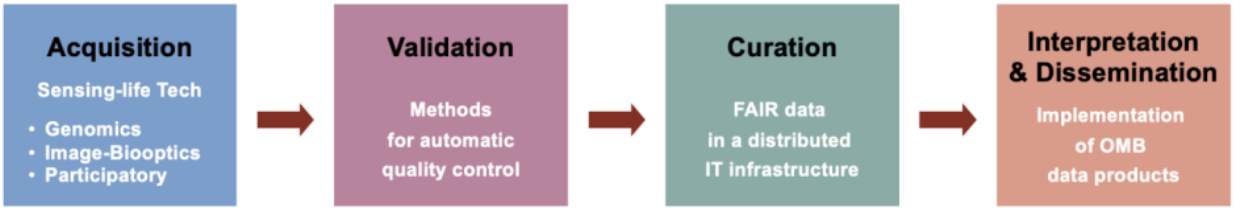

Figure 2.2.1. *The Operational Marine Biology (OMB) information flow, based on an automatic pipeline of information production, from acquisition to interpretation and dissemination.*

OMB data products are currently being developed, with some already confirmed, such as those in WP2 related to Genomics (Bouquieaux, M-C, 2023). These include indicator maps for species diversity, intraspecific genetic variation, and the occurrence of non-indigenous species.

2.2.1 Species diversity map

This is a workflow to showcase diversity over the European observatories for all successive ARMS-MBON deployments and ANERIS generated genomic data (Bouquieaux, M-C, 2023). Output maps capture the *phyla* or lower level taxa detected at the observatories, a map representing the number of species per location and of course, a list of species found per location. See an example in Figure 2.2.1.1 below.

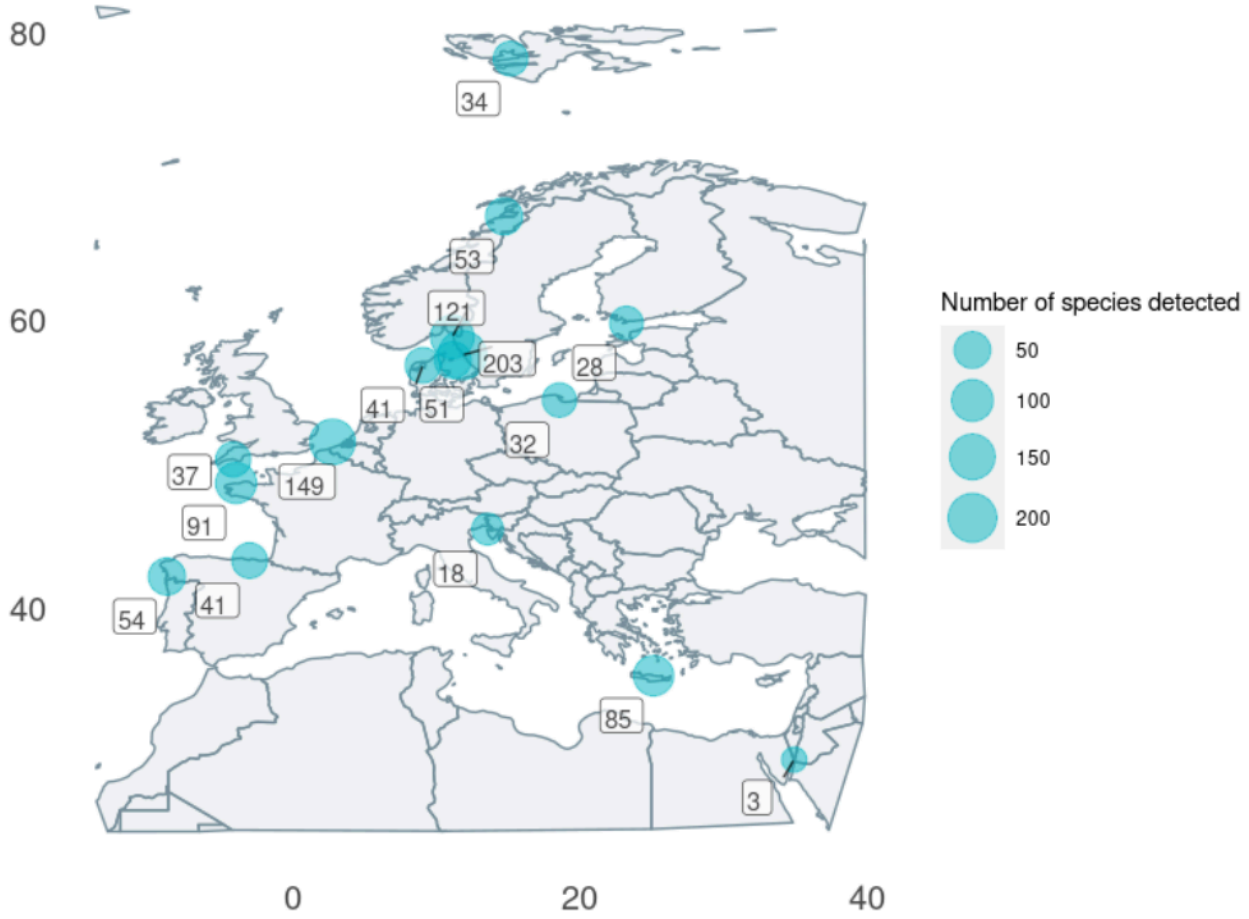

Figure 2.2.1.1. *Number of species detected per ARMS-MBON deployment location.*

2.2.2 Intraspecific genetic variations map

This workflow extracts the sequence of each species found within the ARMS-MBON and ANERIS generated dataset. Sequences from reference databases for each species will also be extracted and added to the original sequencing data. The alignment tool Mafft[17] is then used to create an alignment per species. These alignments are then analyzed to produce haplotypes network shown in Figure 2.2.2.1. Haplotype networks will be represented geographically over Europe and mapped on phylogenetic trees.

[17] Mafft aligner: https://mafft.cbrc.jp/alignment/software/

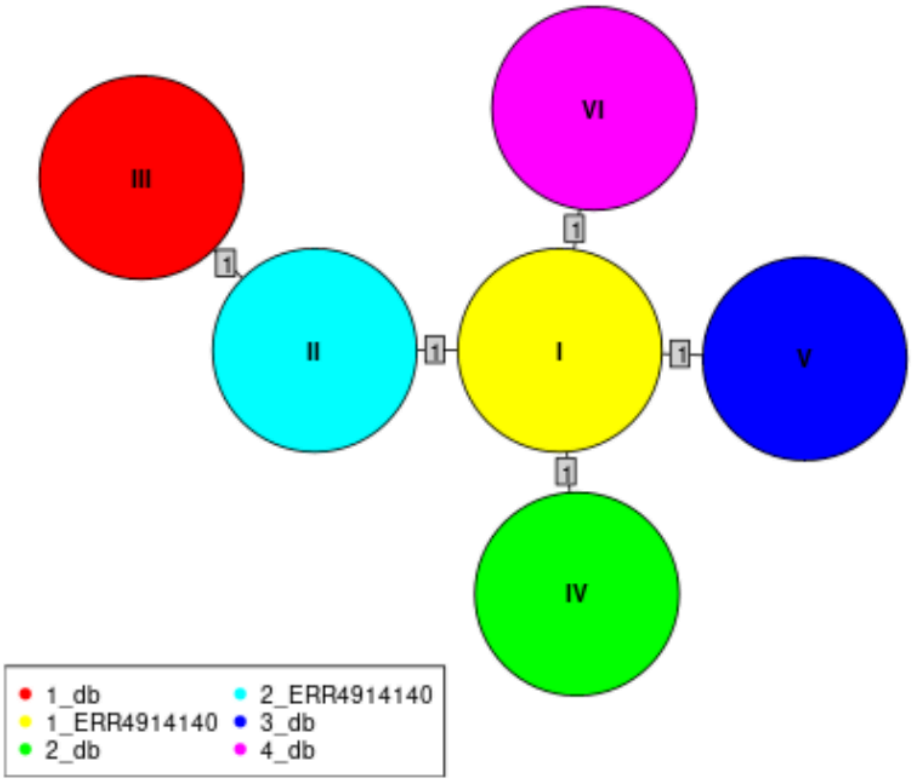

Figure 2.2.2.1. *Haplotype network obtained from ARMS sequence data for the benthic harpacticoid.*

2.2.3 Non-indigenous species map

The output of this workflow consists of a list of the invasive species found per location, and a map representing where those species have been found. The workflow compares the list of the detected species of the samples to a list of invasive species coming from the World Register of Invasive Marine Species (WRIMS[18]) as well as checks for distance to nearest detection of a species using data aggregators like GBIF.

[18] World Register of Invasive Marine Species: https://www.marinespecies.org/introduced/

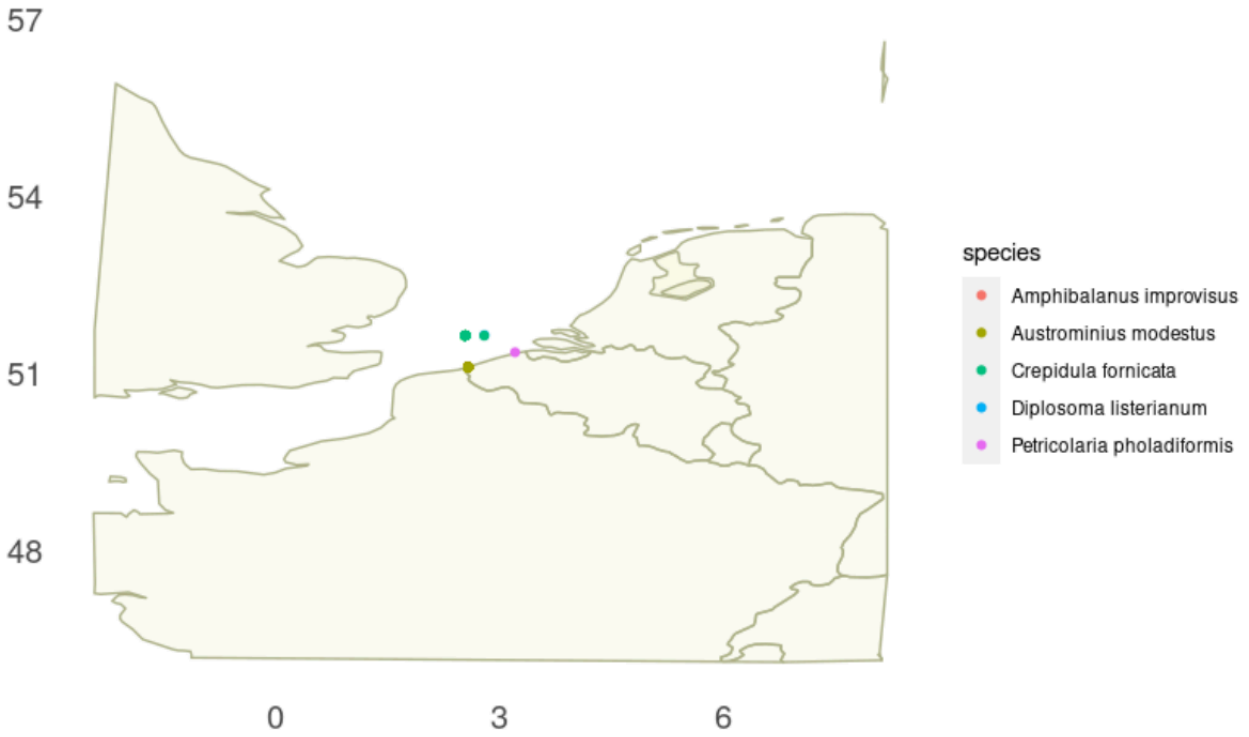

Figure 2.2.3.1. *Non-indigenous species detected on the ARMS deployment located in the Belgian part of the North Sea.*

2.3 Shortlisted ANERIS artifacts

What does it mean to make an artifact available in the EOSC ecosystem? It means that the artifact is available for others to make use of it, either freely or under certain conditions specified by the owner of the artifact. Common examples are project reports (i.e., documents), FAIR data, FAIR software, and services. Most of the artifacts from ANERIS presented in sections 2.1 and 2.2 above (KERs and OMB data products, respectively) can be made available in the EOSC ecosystem (see Table 2.3.1), with the exception of two: the EMUAS technology and the ANERIS cyberinfrastructure. See the discussion in the next paragraphs for more details.

EMUAS is a multi camera system that is set to improve the quality of images acquired in underwater observatories. The general public will have access to the images produced with EMUAS but the technology itself is for internal use to the project. Therefore, there will be an indirect exploitation of the EMUAS technology where general users can benefit from the outputs generated with it, but users will not be able to make a direct use. Likewise, the ANERIS cyberinfrastructure is used to enable end-user services and execute internal workflows but the general public will not be able to access the cyberinfrastructure directly.

Table 2.3.1 below summarizes the list of artifacts produced in the ANERIS project and how the EOSC community will benefit from each of them. Later in Section 3 we detail the available options to make these artifacts available in the EOSC ecosystem.

Table 2.3.1: *Summary of ANERIS artifacts*

ANERIS outputs	Can be shared in EOSC?
ANERIS Technologies	
NANOMICS	Yes, as methodology document
MARGENODAT	Yes, as software
SLIM-2.0	Yes, as software and service
AIES-ZOO	Yes, as software
AIES-PHY	Yes, as software
AIES-MAC	Yes, as software
ATIRES	Yes, as software
EMUAS	Indirectly, via the images produced
AMAMER	Yes, as software
AWIMAR	Yes, as software and service
AMOVALIH	Yes, as software and service
ANERIS Training	
User documentation	Yes, as document
Training material	Yes, as document
Guidelines	Yes, as document
ANERIS Co-Design	
Co-Design methodology	Yes, as document
ANERIS Cyberinfrastructure	
Cyberinfrastructure	Indirectly, via ANERIS services
ANERIS Community	
Community	Yes, join the similar communities
Operational Marine Biology data products	
Species diversity map	Yes: software, service, data source
Intraspecific genetic variations map	Yes: software, service, data source
Non-indigenous species map	Yes: software, service, data source

3. EOSC integration plan

After the introduction to EOSC and its offering for ANERIS in Section 1, and the presentation of ANERIS results that are suitable for EOSC in Section 2, the goal of this Section is to provide an integration plan with steps to make ANERIS results become part of EOSC.

3.1 Software

The easiest way to make your software available in EOSC is via Zenodo. All content available in Zenodo will automatically appear in the EOSC EU Node Resource Hub.

Developers can freely upload the source code of the software along with the corresponding metadata to Zenodo. Another popular choice is to host the source code on GitHub and enable its automatic synchronization with Zenodo. This way every new release of the software on GitHub will create a new entry in Zenodo with a corresponding DOI.

Table 3.1.1 below shows the candidate list of software packages developed in ANERIS that could be uploaded to Zenodo.

Table 3.1.1: *Software packages developed in ANERIS*

Software	Availability
Topic: Genomics	
MARGENODAT	Under development
SLIM-2.0	Under development
Topic: Imaging	
AIES-ZOO	Under development
AIES-PHY	Under development
AIES-MAC	Under development
ATIRES	https://sea-thru-nerf.github.io/
Topic: Citizen Science	
AWIMAR	Under development
AMAMER	Under development
AMOVALIH	Under development
Operational Marine Biology data products	
Species diversity pipeline	Under development
Intraspecific genetic variations pipeline	Under development
Non-indigenous species pipeline	Under development

Software packages uploaded to Zenodo may be included in two Zenodo communities: 1) ANERIS and 2) EU Open Research Repository. The ANERIS community was created by the project to host project outputs (deliverables, presentations, etc.). The EU Open Research Community is managed by CERN on behalf of the European Commission to gather all outputs stemming from Horizon Europe projects.

3.2 Documents (reports, presentations, training material)

As well as software, Zenodo is the easiest way to make documents available in EOSC, including: reports, deliverables, presentations, posters, training material, etc.

The ANERIS project created a Zenodo community to upload the eligible outputs of the project in Zenodo. Table 3.2.1 below shows examples of the entries that are already uploaded to Zenodo.

Table 3.2.1: *Example documents uploaded to the ANERIS community in Zenodo.*

Document	Availability
Topic: Imaging	
Image Classification and Segmentation using Deep Learning	Zenodo DOI
Underwater Image Restoration	Zenodo DOI
Topic: Citizen Science	
Upload data in MINKA - User guide	Zenodo DOI
MINKA APP - User guide	Zenodo DOI

As the time goes by the ANERIS project will continue uploading outputs to the ANERIS community in Zenodo. Moreover, the same content will also be included in the EU Open Research Community managed by CERN on behalf of the European Commission to gather all outputs stemming from Horizon Europe projects.

To illustrate the automatic synchronization between Zenodo and the EOSC EU Node Resource Hub, see how the entry[19] in Zenodo for the Underwater Image Restoration presentation corresponds to the entry[20] in the EOSC EU Node Resource Hub.

3.3 Methodologies

There are two methodologies that will come out of the ANERIS project (Table 3.3.1), in two different areas: 1) NANOMICS (one of the ANERIS technologies) and the co-design methodology (identified as Key Exploitable Result in ANERIS in (Salazar Forn, X., 2023)).

[19] Underwater Image restoration presentation in Zenodo.
[20] Underwater Image restoration presentation in the EOSC EU Node Resource Hub.

Table 3.3.1: *Methodologies coming out of ANERIS*

Methodology	Availability
NANOMICS	Under development
ANERIS co-design methodology	Under development

NANOMICS is focused on creating and evaluating sampling protocols for the different study sites and ecosystems (water samples, intertidal, sediments, etc.) using Oxford Nanopore[21] sequencing. On the other hand, the co-design methodology helps the ANERIS partners with the development and validation of technological solutions in the project.

The suggestion to make these two methodologies available in the EOSC ecosystem is to publish the relevant documents via Zenodo, as explained for Software and Documents in the two previous sections.

3.4 Publications

Instead of Zenodo, the recommendation for ANERIS-centric publications is to use Open Research Europe[22] for preprints or publications. At the time of this writing an open peer review service is currently only available for Horizon 2020 and Horizon Europe beneficiaries.

Open Research Europe is a peer-reviewed publishing platform. Part of its workflow is like that of a preprint server: as a first step authors first rapidly publish their work as a preprint. However, the important distinction is that publication is always automatically followed by invited transparent open peer review. Open Research Europe manages the peer review service, thus completing the publication process. Open Research Europe accepts original research that has not been published before and is not considered for publication elsewhere. Once submitted as a preprint in Open Research Europe, the publications cannot be submitted elsewhere to be peer-reviewed.

Beneficiaries of European Commission funding and staff are free to choose where they wish to publish their research. However, the hope is that the breadth of article types supported, alongside the speed and ease of publication, the rigour and transparency of the peer review process and the high scientific standards for all disciplines coupled with the absence of author fees will be attractive to many researchers.

In addition to traditional research articles, Open Research Europe allows researchers to publish:

- Software Tools Articles[23]: A Software Tool is an article type used to describe novel algorithms, codes, containers, web-apps or workflows and highlights the rationale behind

[21] Nanopore technology: https://nanoporetech.com/platform/technology
[22] Open Research Europe: https://open-research-europe.ec.europa.eu/
[23] Software Tools Articles in Open Research Europe.

the software; examining why it was built, input data sets, and examples of outputs from the tools. The aim is to increase the discoverability of the software itself, allowing experts and researchers to gain the credit they deserve for the software created.

- Data Notes[24]: Data Notes are an article type offered on Open Research Europe which allows researchers to describe their scientific dataset. Data Notes do not include any analyses or conclusions, but promote the discoverability and potential reuse of research data by providing a detailed description of the dataset itself. This gives credit to data producers with a citable, peer-reviewed publication, and supports new research collaborations across disciplines.
- Registered Reports[25]: Receive valuable feedback on your methods and analysis before the data collection stage to enhance your research article and shift the focus of your work away from the results and back to the research question.
- Method Articles[26]: A significant amount of time goes into testing and optimizing a method that researchers cannot fully explain in a traditional Research Article. This is where Method Articles come in. Method Articles allow researchers to provide an in-depth exploration of the method used during the analysis process. This article type also benefits from earlier feedback from the community in the form of open peer review.

Open Research Europe includes a comprehensive guide[27] to support authors to make their publications FAIR. For more information about Open Research Europe, please read its FAQ page[28].

3.5 Data Sources

At the time of this writing the Rules of Participation (RoP) are still to be defined. RoPs will define what Data Sources are eligible to be included in the EOSC EU Node Resource Hub.

Nonetheless, at this point it is still worth identifying the list of candidate Data Sources from ANERIS that could potentially be integrated in EOSC (see Table 3.4.1).

Table 3.4.1: *Candidate Data Sources in ANERIS*

Data Source	Availability
OMB: Species diversity: inventory list of species in a map	Under development
OMB: Intraspecific genetic variations (haplotypes map)	Under development
OMB: Non-indigenous species occurrence in a map	Under development
AWIMAR	Under development

[24] Data Notes in Open Research Europe.
[25] Registered Reports in Open Research Europe.
[26] Method Articles in Open Research Europe.
[27] Open Research Europe guide for FAIR publications.
[28] Open Research Europe FAQ: https://open-research-europe.ec.europa.eu/faqs/

The list of candidate Data Sources shown in Table 3.4.1 includes AWIMAR as a source of images with curated classification from underwater citizen observations, and the currently defined OMB Data Products from WP2: species diversity map, haplotypes map, and non-indigenous species map. Please note that these OMB Data Products will also be an interactive service (a map) but it is expected that they can be queried to extract data about the underlying data used for creating the maps, and that's why they are considered a Data Source as well.

Additional OMB Data Products will be defined by WP3 and will be reviewed to check whether they can also be considered Data Sources for EOSC.

3.6 Services

At the time of this writing the Rules of Participation (RoP) are still to be defined. RoPs will define what Services are eligible to be included in the EOSC EU Node Resource Hub.

Nonetheless, at this point it is still worth identifying the list of candidate Services from ANERIS that could potentially be integrated in EOSC (see Table 3.5.1).

Table 3.5.1: *Candidate Services in ANERIS*

Services	Availability
OMB: Species diversity: inventory list of species in a map	Under development
OMB: Intraspecific genetic variations (haplotypes map)	Under development
OMB: Non-indigenous species occurrence in a map	Under development
AWIMAR	Under development
AMOVALIH	Under development
SLIM-2.0	Under development

The list of candidate Services shown in Table 3.5.1 includes: 1) AWIMAR as the platform to provide images from underwater citizen observations, 2) AMOVALIH a hybrid intelligent system for classifying marine life images, integrating reputation-based classification systems driven by human input with advanced automatic identification systems, 3) SLIM-2.0 as the virtual environment for genomic data analysis, and 4) the interactive maps to be created by the currently defined OMB Data Products from WP2.

Additional OMB Data Products will be defined by WP3 and will be reviewed to check whether they can also be considered Services for EOSC.

3.6.1 EOSC AAI and EGI Check-in

Authentication and Authorization Infrastructure (AAI) in EOSC allows users to use the same (existing) credentials across the EOSC ecosystem.

EOSC AAI is a joint effort among Research Communities and Research Infrastructures that work together to form a federation and enable a Single Sign-On experience across their respective services.

EGI Check-in[29] is an AAI solution that belongs to the EOSC AAI Federation. New services emanating from the ANERIS project can easily align their AAI with EOSC by integrating with EGI Check-in (see Figure 3.6.1.1).

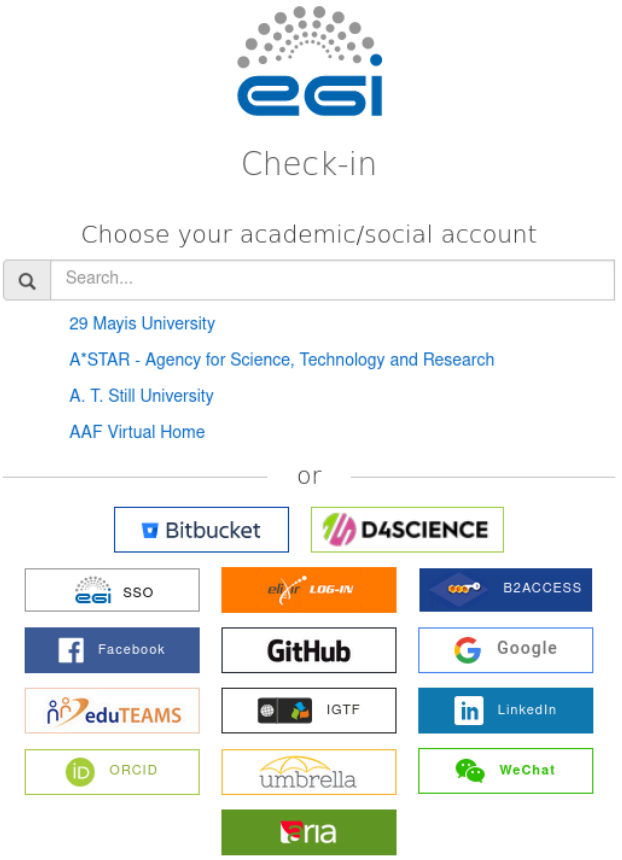

Figure 3.6.1.1. Screenshot showing the EGI Check-in page.

Figure 3.6.1.1 shows the Check-in login page where users are presented with multiple options. In the ideal scenario, users search and find their home institution and reuse their existing credentials. Alternatively, users can also reuse existing credentials in popular social media services such as Facebook, LinkedIn or Google.

A popular choice among researchers is the use of ORCID[30] credentials since it allows accessing the same credentials when moving between different job positions across their career. With EGI Check-in, ORCID is one of the multiple options for the users to simplify access across services.

[29] EGI Check-in: https://www.egi.eu/service/check-in/
[30] ORCID: https://orcid.org/

AWIMAR is one of the first candidate services from ANERIS to potentially benefit from the integration with EGI Check-in. This will be evaluated and revisited in the next iteration of this deliverable due in M36 (December 2025).

3.7 EOSC Nodes

The EOSC ecosystem is envisioned to be a network of interconnected autonomous nodes that collectively provide federated services, such as computing resources, data storage, and data management tools, all operating under a common framework of standards, policies, and best practices. This ensures interoperability and enables researchers to combine resources from different nodes for their research projects. As such, it is expected that more EOSC Nodes will appear. Horizon Europe projects EOSC Beyond[31] and OSCARS[32] intend to help establish national and/or thematic EOSC Nodes.

The EOSC EU Node is the first of the EOSC Federation. However, 1) the minimum requirements of an EOSC Node and 2) the enrollment of eligible Nodes in the EOSC Federation are still being discussed at the time of this writing[33].

Instead of creating a new Node it might be more appropriate for the ANERIS consortium to join an existing Node in the EOSC Federation. The landscape is still uncertain but the ENVRI community, part of the Science Clusters (Petzold, A., 2024), is already following the steps to form a new EOSC Node. Other institutions may follow a similar process but with a more specific focus on Marine Science. The suggestion is to follow the updates on the horizon and join efforts with existing initiatives in order to avoid duplication of effort rather than working on the creation of a new EOSC Node of its own, if possible.

[31] EOSC Beyond project: https://www.eosc-beyond.eu/
[32] OSCARS project: https://science-clusters.eu/
[33] Towards a fully fledged EOSC Federation

Conclusions

Section 1 introduced the European Open Science Cloud (EOSC) and the ways that researchers and contributors can become part of it. Section 2 described the outputs from the ANERIS project that are eligible for onboarding into EOSC. For each type of artifact produced by the ANERIS project, Section 3 suggested the best way to make it available in the EOSC ecosystem, including also the option for the ANERIS consortium to become part of an EOSC Node in the EOSC Federation as part of existing nodes in the domain of Marine Sciences or Environmental Sciences.

Finally in this Section we summarize in Table 4.1 the suggested integration option for each of the artifact types produced in ANERIS.

Table 4.1: *EOSC Integration Plan summarized*

ANERIS outputs	Suggested integration with EOSC
Software	Zenodo
Documents: reports, training material, etc.	Zenodo
Methodologies	Zenodo
Publications	Open Research Europe
Data Sources	EOSC EU Node Resource Hub
Services	EOSC EU Node Resource Hub
Become an EOSC Node	EOSC Federation via existing node

Documents (reports, deliverables, presentations, etc.), methodologies as well as software are best suited to be uploaded in Zenodo since the entries in Zenodo will be automatically added to the EOSC EU Node Resource Hub.

Consortium members are free to choose where they wish to publish their research. However, Open Research Europe was presented as a novel (and free of charge) alternative with an open peer-review process in place that will guarantee alignment with Open Science policies set by the European Commission.

The Data Sources and Services identified in ANERIS are initially suited to become part of the EOSC EU Node Resource Hub. However, the Rules of Participation defining the procedure to do this are still to be defined. At the moment we have just gathered a candidate list and we hope that before the second submission of this deliverable (December 2025) we will be able to provide more details about this onboarding process.

The final option for full integration with the EOSC ecosystem is to create a new EOSC Node to join the EOSC Federation, though the requirements and eligibility criteria are still undetermined. It may be more suitable for the ANERIS consortium to join an existing Node. The landscape is uncertain, but the ENVRI community is already forming a new EOSC Node. Other institutions might do the same with a focus on Marine Science. It is recommended to follow updates and collaborate with existing initiatives to avoid duplicating efforts, if possible.

References

Petzold, A., Hienola, A., Ewbank, J., Tedds, J., Lamanna, G., Bird, I., Gotz, A., Bodera, J., de Jong, F., & Wolff-Boenisch, B. (2024). Science Clusters: Position statement on operational commitment to EOSC and Open Research. Zenodo. https://doi.org/10.5281/zenodo.10732049

Salazar Forn, X., (2023). Exploitation and Sustainability. Deliverable D6.3 EU Horizon Europe ANERIS Project, Grant agreement No. 101094924

Luna-Valero, S., Osychenko, O., Queral, J., Rodero, I. & Sipos, G. (2023). Technology system architecture and integration plan. Deliverable D5.1 EU Horizon Europe ANERIS Project, Grant agreement No. 101094924

Díez, R., Luna-Valero, S. & Rodero, I. (2023). Validated edge-to-cloud deployment for marine biology data lifecycle. Deliverable D5.2 EU Horizon Europe ANERIS Project, Grant agreement No. 101094924

Bouquieaux, M-C., Hablützel, P., Heynderickx, H., Lagaisse, R., Deneudt, K., Díez Lázaro, R., Luna-Valero, S., Rodero, I. (2023). Genomics pipelines and gateways validated on the computational infrastructure through case study CS2. Milestone 15 EU Horizon Europe ANERIS Project, Grant agreement No. 101094924