NEANIAS
Novel EOSC services for Emerging Atmosphere, Underwater and Space Challenges

Whitepaper: Onboarding EOSC hub: best practices and lessons learnt

Jozsef Kovacs (SZTAKI), Gergely Sipos (SZTAKI), Eva Sciacca (INAF), Cristobal Bordiu (INAF), Nikos Chondros (NKUA), Christos Tsiakaliaris (ATHENA), George Papastefanatos (ATHENA), Marco Molinaro (INAF), Daniel Martinez Sala (RICOH).

Disclaimer

NEANIAS is a Research and Innovation Action funded by European Union under Horizon 2020 research and innovation programme, via grant agreement No. 863448.

NEANIAS is project that comprehensively addresses the 'Prototyping New Innovative Services' challenge set out in the 'Roadmap for EOSC' foreseen actions. It drives the co-design, delivery, and integration into EOSC of innovative thematic services, derived from state-of-the-art research assets and practices in three major sectors: underwater research, atmospheric research and space research. In each sector it engages a diverse set of research and business groups, practices, and technologies and will not only address its community-specific needs but will also enable the transition of the respective community to the EOSC concept and Open Science principles. NEANIAS provides its communities with plentiful resource access, collaboration instruments, and interdisciplinary research mechanisms, which will amplify and broaden each community's research and knowledge generation activities. NEANIAS delivers a rich set of services, designed to be flexible and extensible, able to accommodate the needs of communities beyond their original definition and to adapt to neighbouring cases, fostering reproducibility and re-usability. NEANIAS identifies promising, cutting-edge business cases across several user communities and lays out several concrete exploitation opportunities.

This document has been produced receiving funding from the European Commission. The content of this document is a product of the NEANIAS project Consortium and it does not necessarily reflect the opinion of the European Commission. The editor, author, contributors and reviewers of this document have taken any available measure in order for its content to be accurate and lawful. However, neither the project consortium as a whole nor the individual partners that implicitly or explicitly participated in the creation and publication of this document may be held responsible for any damage, financial or other loss or any other issue that may arise as a result of using the content of this document or any of the project outputs that this document may refer to.

The European Union (EU) was established in accordance with the Treaty on the European Union (Maastricht). There are currently 28 member states of the European Union. It is based on the European Communities and the member states' cooperation in the fields of Common Foreign and Security Policy and Justice and Home Affairs. The five main institutions of the European Union are the European Parliament, the Council of Ministers, the European Commission, the Court of Justice, and the Court of Auditors (http://europa.eu.int/).

Table of Contents

Abstract

The NEANIAS EU project develops and integrates services to EOSC in the field of Underwater, Atmospheric and Space research. The objective is to share these thematic services with scientific researchers of EOSC communities on these fields and beyond. The project completed the development and EOSC onboarding of 4 of its thematic services: SPACE-VIS ViaLactea, UW-MAP, ATMO-FLUD and SPACE-ML CAESAR. This white paper reports on the EOSC onboarding experience, i.e., on the registration of these services in the EOSC Portal and summarises the lessons learnt during the process.

1.   Introduction

The NEANIAS EU project develops and integrates services to EOSC in the field of Underwater, Atmospheric and Space research. The objective is to share these thematic services with scientific researchers of EOSC communities on these fields and beyond.

At the time of writing this whitepaper, the project completed the development and EOSC onboarding of 4 of its thematic services: SPACE-VIS ViaLactea Service [2], UW-MAP [3], ATMO-FLUD [4], SPACE-ML CAESAR service [5]. This whitepaper reports on the EOSC onboarding experience, i.e., on the registration of these services in the EOSC Portal [6].

Section 2 starts with a short overview of EOSC in order to give an insight to the initiative, the landscape. In Section 3 the EOSC onboarding process is described and the NEANIAS implementation is summarised. Section 4 provides details about the onboarding experiences of the 4 NEANIAS thematic services, while Section 5 pulls together the main lessons learnt from the activity.

2.   Overview of European Open Science Cloud (EOSC)

The European Open Science Cloud (EOSC) is a European Commission initiative aiming at a trusted, virtual, federated environment in Europe to store, share and re-use digital output from research (publications, data and software) across borders and scientific disciplines. The envisaged infrastructure is established by aggregating services, software, data and other types of scientific outputs from a diverse set of providers.

The EOSC initiative started in 2015, and the first phase of EOSC development recently finished under a multi-layered, interim governance structure that was active from November 2018 until the end of 2020. In the initial phase of development until the end of 2020, the Commission invested around €320 million to start prototyping the EOSC through project calls in Horizon 2020 - the Commission's research and innovation funding programme.

EOSC has a "core", the set of services providing the means to discover, share, access and re-use data and services. The initial version of the EOSC core was developed and operated by the EOSC-hub project between 2018-2020. Onboarding services to EOSC therefore used to be called 'onboarding to the EOSC hub', and that is what motivated the title of this whitepaper (which was defined during the running of EOSC-hub). The EOSC Portal (https://eosc-portal.eu/) is part of the EOSC core, and since 2021 it is operated and further developed by

the EOSC-Enhance project. The EOSC Portal provides a European delivery channel connecting the demand-side and the supply-side of the EOSC and all its stakeholders.

EU countries and countries associated with Horizon 2020, represented in the EOSC Governance Board, agreed unanimously to run the EOSC as a co-programmed European Partnership under Horizon Europe from 2021. Horizon Europe is the Commission's research and innovation funding programme, succeeding Horizon 2020 from 2021. The new governance model agreed with EU countries for the next EOSC implementation phase after 2020 will be tripartite including:

- The EU represented by the Commission
- The European research community represented by the EOSC Association
- EU countries and countries associated with Horizon Europe represented through a Steering Board to be set up in 2021 outside of the EOSC Association

3. Onboarding NEANIAS services in EOSC

The heart of EOSC (aka. EOSC core) is an integration and management system that enables the delivery of services, software and data from e-infrastructures, major research infrastructures and a growing number of other providers. This core builds on mature processes, policies and tools from the leading European federated e-Infrastructures to cover the whole life cycle of services, from planning to delivery. The core enables 'EOSC exchange' the catalogue of services shared via EOSC by the providers.

Both EOSC core and EOSC exchange are exposed to the public (both providers, users and supporters) via the EOSC Portal which acts as a single contact point to discover, access, use and reuse a broad spectrum of resources for advanced data-driven research. Through the virtual access mechanism, more scientific communities and users have access to services supporting their scientific discovery and collaboration across disciplinary and geographical boundaries. EOSC poses the following requirements, aka. 'Rules of Participation' (RoP), on services:

- The service is accessible to users outside its original community.
- The service is described through a common template focused on value proposition and functional capabilities.
- At least one service instance is running in a production environment available to the user community.
- Publish Research data which is Findable, Accessible, Interoperable and Reusable.
- Release notes and sufficient documentation are available.
- Helpdesk channels are available for support, bug reporting and requirements gathering.

Our EOSC integration plan was defined in Q4 2019 (NEANIAS D8.1 [1]), based on the capabilities and possibilities of the EOSC core and EOSC Portal at that time. While some of the onboarding tools evolved since 2019, the onboarding process fundamentally remained the same, with the following minimal steps:

A. Provider representative registers him/herself into the EOSC Portal.
B. Provider onboards (and updates) the Provider information.

C. Provider onboards (and updates) the Resources offered by the Provider.

The forms that have to be filled during these steps ensure compliance with the above mentioned RoP as it captures relevant aspects of the provided services (e.g., description in a common template, pointer to running instance, link to helpdesk) and the submitted data is checked manually by the EOSC onboarding team. If the three-phase onboarding process is successful, then the Provider is registered to the EOSC Portal and their Resources become publicly accessible on its web interface at [6].

Additional, optional integration steps with the EOSC core are also possible after/during the onboarding process. These steps were explained in NEANIAS Deliverable D8.1. These optional steps are:

- Integration with EOSC AAI
- Integration with EOSC service availability-reliability monitor
- Integration with EOSC Usage accounting
- Integration with the EOSC Helpdesk
- Alignment with the EOSC Service Management System (SMS).

Table 1 provides a summary of the steps that were followed by the 3 NEANIAS thematic services which are now visible in EOSC at [2][3][4].

Table 1: EOSC integration status of NEANIAS thematic services in July 2021.

	SPACE-VIS ViaLactea	UW-MAP	ATMO-FLUD	SPACE-ML CAESAR
Onboarding in EOSC Portal	Completed [2]	Completed [3]	Completed [4]	Completed [5]
Integration with AAI	Not relevant because the services are open access, but login is required through the NEANIAS AAI.			
Integration with Monitor	The services are monitored by the NEANIAS Monitor service instead of the EOSC Monitor.			
Integration with Accounting	The services are integrated to the NEANIAS Accounting system, but not to the EOSC (EGI) Accounting.			
Integration with Helpdesk	Not relevant because the services serve users via the NEANIAS Helpdesk.			
Alignment with SMS	Completed, via the NEANIAS SMS which is based on FitSM, just like the EOSC SMS.			

4. Onboarding experiences in NEANIAS

4.1. SPACE-VIS ViaLactea and SPACE-ML CAESAR Services

Both ViaLactea and CAESAR services are provided by the Istituto Nazionale di Astrofisica (INAF) and have been recently published in the EOSC portal.
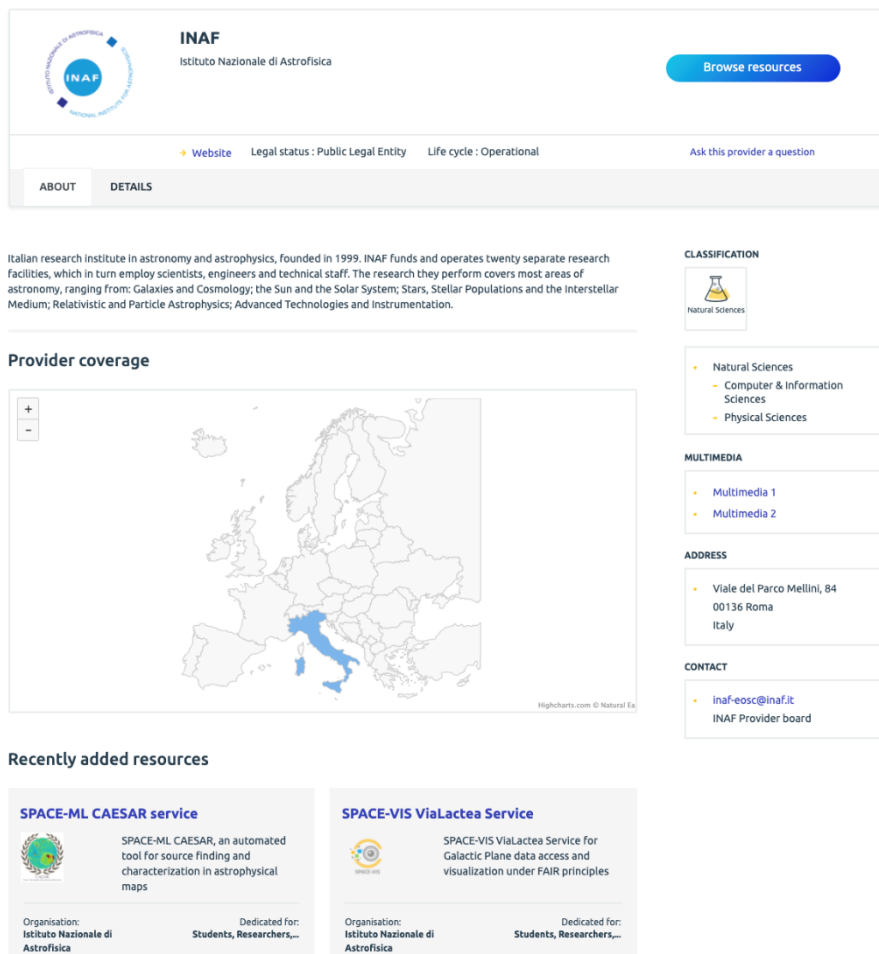
The SPACE-VIS ViaLactea Service allows for studying star formation phenomena in our galaxy, the Milky Way. In NEANIAS, the service has been improved by exploiting the cloud infrastructures toward supporting the FAIR principles and applying the standards of the Virtual Observatory. In particular the ViaLactea data service resources have been deployed on the GARR Cloud Platform and was integrated with the NEANIAS AAI authorization functionalities to protect the surveys according to each observatory data policy. The service is currently accessed by a Visual Analytics desktop application for the searching, access, visualization and analysis of the most updated surveys of the Galactic Plane, from the infrared to the radio band.

Likewise, CAESAR (Compact and Extended Source Automated Recognition) service is a source finder application tailored to the requirements of next generation telescopes, such as the Square Kilometer Array. The amount of produced data will need automated source finding techniques. CAESAR is able to efficiently extract and parameterize both compact and extended sources from large radio maps, and it is being currently to process data from the SKA pathfinders: the ASKAP telescope in Australia and the Meerkat array in South Africa. In NEANIAS, it has been extended to be used by REST-APIs from containers (both Docker and Singularity) and deployed on the GARR Cloud Platform Service and on the Container Service with a user-friendly User Interface using the NEANIAS templates.

For the EOSC onboarding process it was firstly necessary to identify the Authorized Representative of INAF as new EOSC Provider, and this process needed some time to discuss internally at INAF and identify the most suitable person. Then the Authorized and Authenticated Representative of INAF onboarded INAF as new EOSC Provider (see Figure 1). We had already collected all the needed information at the NEANIAS Service Catalogue core service, therefore we copied them to the EOSC service registration forms. However, during this process we had some problems regarding required values in form of URLs and we contacted the EOSC Portal help desk, experiencing very slow replies.

Finally, before onboarding the services as new EOSC Resources we finalized the integration with the NEANIAS SMS, that in the case of the ViaLactea service is required for the users to request access to it thus assigning the proper authorization for the survey and data access. Also, we provided the needed legal information regarding the "Terms of Use" and "Privacy Policy" since even if those are not mandatory at this stage, will became soon. These documentations were provided thanks to general template prepared by the NEANIAS team easily customizable according to the features of each service.

Figure 1: INAF EOSC Provider in the EOSC Portal and its services SPACE-VIS ViaLactea and SPACE-ML CAESAR as EOSC Resources.

In the next future, thanks to the collaboration with the University of Portsmouth (UoP) we would like to integrate within the ViaLactea service a simplified web version of the User Interface for the user community to access a reduced set of functionalities of the ViaLactea system. Simultaneously, we are working in close cooperation with the core artificial intelligence team, mainly with the University of Milano Bicocca (UniMIB) and SZTAKI, to integrate deep convolutional neural networks that allows for both source detection and classification.

Therefore, regarding the EOSC Resources future updates, we will need to decide on the best representation of multiple providers: INAF and UoP for the ViaLactea web UI service module, and INAF, UniMIB and SZTAKI for the CAESAR DeepLearning services and modules.

4.2. ATMO-FLUD and UW-MAP services

Both ATMO-FLUD and UW-MAP services are provided by the ATHENA Research Center and have been published in the EOSC portal.

The UW-MAP service (https://uw-map.neanias.eu) is part of the underwater thematic services and regards seabed classification based on multi-spectral multi-beam echosounder data. Most information required for onboarding the service to EOSC was already available from the NEANIAS service catalogue. UW-MAP was the first service from ATHENA submitted to the EOSC Portal Onboarding Team (EPOT) for review and publishing. The feedback received was positive overall, however it was requested that the service provider fills the sections regarding the Terms of Use and Privacy Policy of the service, as this information will become mandatory by Autumn 2021. This information has been provided at https://uw-map.neanias.eu/tos and https://uw-map.neanias.eu/privacy, respectively. Additionally, the EPOT representative expressed some concerns regarding the clarity of the user login process using the NEANIAS Identity Provider which were suitably addressed by providing more detailed instructions both in the NEANIAS Identity Provider landing page and the EOSC Service Description section. After all previous issues were addressed, the EPOT representative authorized the service publishing on the EOSC portal.

The ATMO-FLUD service (https://atmo-flud.neanias.eu/), which is one of the NEANIAS thematic services belonging to the atmospheric research sector, was onboarded to EOSC in June 2021. The first action for onboarding was to ensure that ATHENA is registered as a service provider at EOSC, a process that required ATHENA to designate an authorized representative for EOSC.

Thanks to the NEANIAS Service Catalogue core service, we were already accustomed with the service onboarding process and the required inputs that we needed to provide to EOSC. In fact, we had already registered, at an earlier stage, in the NEANIAS service catalogue most of the information required by EOSC, and we simply verified and copied them to the EOSC service registration forms. Additionally, thanks to the fact that our institution (ATHENA) had already onboarded the UW-MAP service, and we had to address minor issues at that point (e.g., regarding user authentication), ATMO-FLUD's onboarding was smooth and straightforward.

However, for the ATMO-FLUD service we were very concerned about the Terms of Use, as we require attribution for all outputs of the service that are used in scientific publications. To this end, we were involved in all discussions with the legal team that shaped the "Terms and Conditions" document that we were provided as a template to all NEANIAS services and we adapted for our service (accessible here: https://atmo-flud.neanias.eu/Terms and Conditions). Even though the Terms and Conditions and the Privacy Policy documents were not yet mandatory information during the EOSC registration process, for our own reasons we wanted our service to be legally well protected before we "advertise" it at the EOSC Marketplace. In fact, our legal worries were the only reason the service was onboarded in June 2021 and not earlier.

5.   Conclusion - Lessons learnt

5.1. Positive overall experience

After having onboarded four (4) services, originating from two (2) different providers, on EOSC, we found the onboarding process quite straight forward and not difficult to use. Of course, as one would normally expect, several minor issues were encountered during the onboarding process, which are briefly discussed below. Some of the issues belong to the

service providers' area of responsibility, while some other ones are related to EOSC procedures.

5.2. Identification of the Authorized Representative of Provider

Although this should be something very easy to address, both service providers of the four services faced some difficulties and delays in identifying and authorizing a person to act as the "Representative of the Provider" in what EOSC is concerned. To our understanding until now, this is due to the type (research institutes), the organization and the internal procedures of the two providers. This resulted to delaying the beginning of the onboarding process, as the registration (and authorization) of this person, i.e., the Authorized Representative of the Provider, is a prerequisite phase (Phase A) of the onboarding process.

In order to avoid similar delays in the future, we have advised, under WP8 facilitation, all service providers to identify their EOSC representatives as soon as possible, even way before they start onboarding their services on EOSC.

5.3. Registration of the first resource (service) on EOSC

NEANIAS has developed and populated its own service catalogue before EOSC onboarding. All NEANIAS services are presented there following a service model that adopts the latest EOSC one, namely EOSC Profiles (v3.0). We also have developed the relevant catalogue API for facilitating the integration with EOSC. However, due to restrictions of the EOSC onboarding process, this API cannot be used for the initial registration of a service provider to EOSC or the onboarding of the first service of this provider. This resulted on having to manually "copy-paste" the provider and the service descriptive information from the NEANIAS service catalogue to the EOSC portal.

Exploiting the features of the EOSC available APIs, we are planning the integration of the NEANIAS Catalogue with the EOSC Portal in order to synchronise the information about NEANIAS services between them; of course, this can be done only after the initial (manual) registration of a provider and its first resource on EOSC.

5.4. Terms of Use - Privacy Policy - Access Policy

Although these documents are not currently mandatory for the onboarding of services to EOSC, we chose to prepare them for all four services, adopting relevant templates prepared in collaboration with legal persons, experienced in the subject.

5.5. Support of a service by multiple service providers

In some cases, a service is provided by a service provider but is also supported by another one (or it may even be jointly offered by multiple providers). In order to "link" all providers on EOSC portal to a service, it is necessary for all providers to be registered on EOSC portal. Although reasonable, in order not to delay the onboarding of services having this issue, we chose to ignore the association of the 2nd provider with a service at this stage, and proceeded with onboarding the service. When the 2nd provider gets registered on EOSC portal in the future, we will revisit the service description and associate the 2nd provider with the service.

5.6. Issues during validation

Some minor issues were encountered during the validation of the services information by EOSC Portal Teams (Onboarding Team and Quality Team). These issues, as reported by

NEANIAS service providers, indicatively include the use of personal e-mail accounts as helpdesk/security contacts, some fields requiring values in form of URLs, and issues regarding user authentication. However, these issues were addressed quite easily, after receiving the respective input by EOSC Portal Teams.

References

[1] NEANIAS D8.1 - EOSC Integration Plan.

[2] SPACE-VIS ViaLactea Service: https://marketplace.eosc-portal.eu/services/space-vis-vialactea-service

[3] UW-MAP Service: https://marketplace.eosc-portal.eu/services/uw-map

[4] ATMO-FLUD service: https://marketplace.eosc-portal.eu/services/atmo-flud

[5] SPACE-ML CAESAR service: https://marketplace.eosc-portal.eu/services/space-ml-caesar-service

[6] EOSC Portal: https://eosc-portal.eu

Glossary

AAI – Authentication Authorisation Infrastructure: Set of services that enable users to login and being authorised for access within online services.

ASKAP - Australian Square Kilometre Array Pathfinder (ASKAP) is a radio telescope situated about 800 km north of Perth, in the Murchison region of Western Australia.

EOSC – European Open Science Cloud

EOSC Association – The legal entity that was established as a legal entity on 29th July 2020

EOSC Portal – The graphical, web-based access layer of EOSC to progress the EOSC partnership.

FitSM - A free and lightweight standards family aimed at enabling effective IT service management in the broadest range of organisations.

Onboarding – Registering a service in the EOSC Portal

GARR - Consortium GARR (GARR) is a non-profit association founded under the auspices of the Ministry of Education, University and Research. GARR is provider of cloud and other generic services in NEANIAS.

RoP – Rules of Participation: The set of conditions that services must meet to be acceptable in EOSC.

SMS – Service Management System: Overall management system that controls and supports management of services within an organisation or federation. The SMS can be regarded as the entirety of interconnected policies, processes, procedures, roles, agreements, plans, related resources and other elements needed and used by a service provider to effectively manage the delivery of services to customers.

SPACE-VIS ViaLactea - the service provides an advanced operational solution for data management and visualization of astrophysics FAIR data surveys of the Galactic Plane to study the star formation process of the Milky Way.

URL - Uniform Resource Locator: colloquially termed a web address, is a reference to a web resource that specifies its location on a computer network and a mechanism for retrieving it.

UW-MAP - A Seabed Classification from Multispectral, Multibeam Data service that delivers a user-friendly cloud-based solution integrating cutting-edge machine learning frameworks for mapping several seabed classes, validated for archeological, geo-hazards, energy, and other applications.