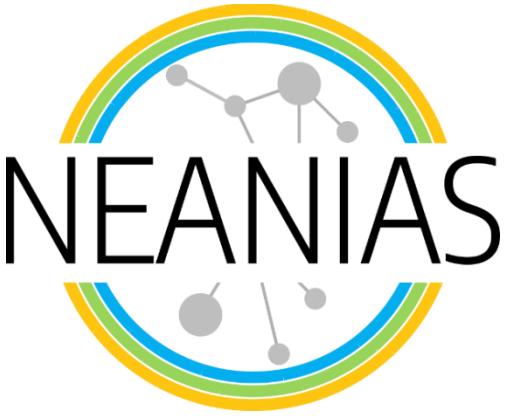

NEANIAS
Novel EOSC services for Emerging Atmosphere, Underwater and Space Challenges

Deliverable D8.1

Deliverable: D8.1 EOSC integration plan

31/07/2020

funded by European Union under Horizon 2020 research and innovation via grant agreement No. 863448

EOSC integration planEOSC integration plan

Document Info

Project Information			
Acronym	NEANIAS		
Name	Novel EOSC Services for Emerging Atmosphere, Underwater & Space Challenges		
Start Date	1 Nov 2019	End Date	31 Oct 2022
Program	H2020-EU.1.4.1.3. - Development, deployment and operation of ICT-based e-infrastructures		
Call ID	H2020-INFRAEOSC-2018-2020	Topic	H2020-INFRAEOSC-2019-1
Grant No	863448	Instrument	RIA

Document Information			
Deliverable No	D8.1		
Deliverable Title	EOSC integration plan		
Due Date	20-07-2020	Delivery Date	31-07-2020
Lead Beneficiary	SZTAKI		
Beneficiaries (part.)	SZTAKI, NKUA, ATHENA, GARR		
Editor(s)	József Kovács (SZTAKI)		
Authors (s)	József Kovács (SZTAKI), Attila Farkas (SZTAKI), Gergely Sipos (SZTAKI), Róbert Lovas (SZTAKI), Claudio Pisa (GARR), George Papastefanatos (ATHENA), Christina Peraki (ATHENA), Nikos Chondros (NKUA)		
Contributor (s)			
Reviewer(s)	Georgios Kakaletris (CITE)		
Workpackage No	WP8		
Version	V0.3	Stage	In progress
Version details	Revision: *985* . Last save: *2020-08-12 , 07:59* Pages: *37* . Characters: *71.720*		
Distribution	Public	Type	Report
Keywords	EOSC, service, integration, plan		

Change Record

Version	Date	Change Description	Editor	Change Location (page/section)
0.1	13/03/2020	Initial TOC	Jozsef Kovacs	All
0.2	03/08/2020	Full draft for internal review	Jozsef Kovacs	All
0.3	05/08/2020	Updated draft based on review feedback	Jozsef Kovacs	All

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 neighboring 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

EOSC integration planEOSC integration plan

Abstract

NEANIAS develops thematic services for EOSC communities on the field of Underwater, Atmospheric and Space research. The project will integrate the thematic services into the EOSC ecosystem after they reach the maturity required by EOSC. This deliverable collects and introduces the integration approaches, procedures, relevant services discovered as part of the EOSC landscape. Based on this information the document then proposes an integration strategy and roadmap.

1. Introduction

The NEANIAS EU project aims to develop and integrate 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.

At the time of writing this deliverable, the project has identified the thematic services and are preparing them for integration. This deliverable aims to provide guidelines for the integration work. The service on-boarding process will be implemented by workpackage WP7 later and this deliverable summarises the most important background knowledge for this work. This knowledge includes key steps to perform, the relevant services of EOSC to integrate to and the strategy of the overall integration procedure.

The structure of this document is as follows: Section 2 starts with a short overview of EOSC in order to give an insight to the initiative and its landscape. In Section 3, the integration guidelines are summarised including the different integration approaches, the key components (e.g. EOSC portal and marketplace), federated services, data management and service management system. The NEANIAS services to be integrated are introduced in Section 4. Section 5 overviews the available EOSC services which are key to the successful integration. Section 6 introduces the integration plan (roadmap) and finally the conclusions are presented in Section 7.

This deliverable has been written based on the current EOSC landscape which is a continuously evolving environment. The next version of this integration plan will reflect to the changes happening in the EOSC landscape until M22 in relation to several fields, such as architecture, interfaces, access mechanism, federated services and so on.

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, with the first version of the European Open Science Cloud officially launched in November 2018, with the EOSC Portal entry point (https://eosc-portal.eu/). The development of EOSC is currently governed by three bodies, as defined by the European Commission Staff Working Document Implementation Roadmap for the European Open Science Cloud:

- The Executive Board, a body tasked to ensure implementation and accountability.
- The Governance Board of EOSC is an institutional group gathering representatives from the Member States and Associated Countries and from the Commission to ensure effective supervision of the EOSC implementation.
- The Stakeholder Forum is the community actively contributing and participating to the European Open Science Cloud (EOSC). It is composed by organisations, projects and initiatives fully committed to support the EOSC vision.

The European Commission is providing financial support to implement the EOSC by means of projects under the EU Framework Programme for Research and Innovation (Horizon 2020). The number of Horizon 2020 projects contributing to EOSC reached over 40 by 2020[1].

Two of the currently running EOSC projects, EOSC-hub and OpenAIRE-Advance, in collaboration with other EOSC projects of the past (EOSCpilot, eInfraCentral) established the central services of EOSC by November 2018, and operate and continuously evolve these since then.

The heart of EOSC is an integration and management system that delivers a catalogue of services, software and data from the EGI Federation, EUDAT CDI, INDIGO-DataCloud, major research infrastructures and a growing number of other providers. This integration and management system (the Hub) 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 Hub aggregates services from local, regional and national e-Infrastructures in Europe, Africa, Asia, Canada and South America.

[1] EOSC Projects: https://eosc-portal.eu/about/eosc-projects

EOSC integration planEOSC integration plan

The Hub is 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.

Our EOSC integration plan is based on the current capabilities and possibilities of the EOSC Portal and its back-end Hub. However, we are aware that the EOSC landscape is evolving:

- The 'EOSC Core' is under definition by the EOSC Architecture Working Group, with a publication expected at the end of 2020. The EOSC Core will be the simplest set of services and interfaces that are required for EOSC to function. We expect that the EOSC Portal, Marketplace, and those federation services that are considered for use (see Section 6) will remain to be part of the EOSC Core, and therefore will stay relevant for NEANIAS.
- A Horizon2020 call, titled '*INFRAEOSC-03-2020 Integration and consolidation of the existing pan-European access mechanism to public research infrastructures and commercial services through the EOSC Portal*' was closed on June 18, 2020. The project to be funded in this call will operate and further develop the EOSC Core (the federation services we consider in Section 6). The project is expected to start in 2021 and run for 30 months.
- 6 Horizon2020 calls, titled *'INFRAEOSC-07-2020 Increasing the service offer of the EOSC Portal'* was closed on June 18, 2020. The projects that will be funded from this call will bring into EOSC (or continue to operate in EOSC) a diverse set of services from the following areas:
    - Distributed and cloud computing resources
    - Data services
    - Services supporting scholarly communication and open access
    - Above the net services (range from simple tools to complex collaborative platforms, including real-time communications and media)
    - Services and resources from non-research public sector data providers
    - Additional research enabling services (text and data mining or Copernicus services)

The projects are expected to start in 2021 and run for 30 months. The projects are expected to cover, to a large extent research management services that are considered for integration in Section 6, or additional services that can be considered in the future (this deliverable will have an update as D8.3 in month 22).

3. Integration guidelines for EOSC

This section is written based on the integration guidelines that exist on the EOSC Portal at July 2020. These guidelines have been provided by the EOSC-hub project in the 'EOSC-hub Integration Handbook for Service Providers' document[2]. The document explains the compulsory and the optional steps that service providers need to take to become service operators in EOSC.

3.1. Overview of integration approaches

The minimum level of EOSC integration is that the service entry should be published in the EOSC Portal and Marketplace. Publication of a service entry ensures that the service meets minimum requirements, it is visible, properly described and accessible for new users. If the service requires users to apply for access, then such user access requests will be submitted via the EOSC Marketplace, then forwarded to the service provider who can evaluate and respond to the requests using the EOSC tools. The service publication is elaborated in Section 3.2.

Additional integration options also exist besides the portal publication, but all those integrations are optional. A provider is free to choose from these additional integration options according to the value seen for the users from these additional ties with EOSC. These additional options are described in Section 3.3, 3.4 and 3.5.

Our intended use of the compulsory publication step, as well as the optional steps are described in Section 6.

3.2. Registering in the EOSC portal

EOSC services can be accessed by users via the EOSC Portal. Any service that wish to be present in EOSC need to apply for inclusion in the Portal via a dedicated form. This application form gathers initial information about the service, reasons for wanting to become an EOSC provider, access conditions and provider contacts. The information is initially reviewed before a Service Description Template (SDT) is created for the requester to fill it in. The SDT covers more detailed information which enables the service onboarding team to determine that the service fulfils all requirements of the 'Rules of Participation' in addition to information needed to populate the EOSC Service Portfolio entry corresponding to the new service. The SDT is a longer form, covering areas such as detailed description of the service, licenses, helpdesk contact, accessibility, etc. The onboarding team also provides customisable templates for some of the items, e.g. for 'acceptable use policy'. These templates can help providers fill any gaps they are currently missing from a complete service description.

[2] EOSC-hub integration handbook for service providers: https://eosc-portal.eu/for-providers

EOSC integration planEOSC integration plan

The submitted information is reviewed for suitability by the onboarding team, then a draft, hidden page is setup about the service within the EOSC Portal. This draft is sent back to the submitter for final validation. After successful validation the service entry is made public and accessible in the EOSC Portal and Marketplace.

3.2.1. EOSC fundamental service requirements

At the time of writing, there are the following fundamental requirements for a service to be onboarded in EOSC:

1. The service falls within the remit of the EOSC activities, i.e. it brings value to users and facilitates them to implement Open Science.
2. It is either an online service (e.g. a web application portal, a web service) or a 'human' service, such as training and consultancy. (plain datasets and software artefacts should not be directly onboarded to EOSC. There are other ways to do that.)
3. The service is mature, reaching 'Technology Readiness Level 7 (TRL7). TRL7 services are 'System prototype demonstration in operational environment', practically meaning that they have been already used by early adopter scientists.
4. The compulsory fields of the service description template are filled during onboarding.

3.3. Integrating with federation services

EOSC-hub offers a set of services that service providers can use to enhance their services from the operational perspective. These services can, for example, simplify how users access the service via federated authentication, they can improve service reliability, they can provide details on capacity consumption by the users, or can simplify user interaction via a helpdesk. This section provides an overview of these services, using the same structure for each.

3.3.1. Federated user authentication and authorisation

The EOSC-hub Authentication and Authorisation Infrastructure (AAI) [28] enables seamless, authenticated access to services and research data in EOSC. The EOSC-hub AAI enables service providers to control access to their services from users holding identities (usernames and passwords) from a very broad set of academic, community or social Identity Providers (IdPs). The EOSC-hub AAI brings together these IdPs, the EOSC-hub service providers (SPs) and intermediary identity management proxies into a single, interoperable infrastructure. The EOSC-hub AAI builds on open technologies including SAML 2.0, OpenID Connect, OAuth 2.0 and X.509v3 to offer a flexible framework for access management. The EOSC-hub AAI comprises different, compatible endpoint services, namely B2ACCESS, Check-in, eduTEAMS and INDIGO-IAM that service providers are free to choose from depending on their preference of technology or provider. The infrastructure also includes Perun, a component that can be used for managing users' access rights to the services, and a set of Token Translation Services to translate between different protocols or technologies when passing identities and user roles to services.

3.3.2. Availability and reliability monitoring

Monitoring is the key service needed to gain insights into an infrastructure. It needs to be continuous and on-demand to quickly detect, correlate, and analyse data for a fast reaction to anomalous behaviour. The challenge of this type of monitoring is how to quickly identify and correlate problems before they affect end-users and ultimately the productivity of their organizations. The features of a monitoring system are monitoring of services, reporting availability and reliability, visualization of the services status, providing dashboard interfaces and sending real-time alerts. Management teams, administrators, service owners can monitor the availability and reliability of the services from a high-level view down to individual system metrics and monitor the conformance of multiple SLAs.

EOSC-hub provides a service monitoring service based on the ARGO system [19]. This ARGO Service collects status results from one or more monitoring engine(s) and delivers status results and/or monthly availability (A) and reliability (R) results of distributed services. Both status results and A/R metrics are presented through a Web UI, with the ability for a user to drill-down from the availability of a site to individual services, to individual test results that contributed to the computed figure. Argo is capable also to send notifications to the service admins in case of a failure/warning on one of the services monitored.

3.3.3. Usage accounting

EGI provides an accounting service [10] for EOSC. This Accounting service can collect, store, aggregate, and display usage information about the following types of services:

- High Throughput Compute
- Infrastructure-as-a-Service cloud virtual machines
- Storage space providers
- Data set providers

This usage data is collected from those Resource Centres that provide the above types of services, and that connect their service endpoints to the centrally managed Accounting Service. Accounting information is gathered from the service by probes and sensors according to certain data formats.

Probes and sensors are deployed locally at the service providers. Data is forwarded from the sensors into a central Accounting Repository where those data are processed to generate various summaries and views for display in the Accounting Portal. Depending on the complexity of the provider the accounting data may go via intermediate repositories that collate accounting data for particular regions, sub-infrastructures or communities. EOSC service providers can either directly publish accounting information into the EOSC Accounting Repository or can do so via an intermediate repository that serves for example a specific region or group of providers. It is up to the provider (group) to use the central repository directly, or to apply an intermediary accounting infrastructure and connect it to EOSC.

3.3.4. Helpdesk

The EOSC-hub Helpdesk [13] is the entry point and ticketing system/request tracker for issues concerning EOSC services. New service providers of EOSC can integrate into the Helpdesk and this results in:

- a corresponding support topic listed on the Helpdesk user interface (for users to ask questions or raise issues directly to the provider)
- the provider support team to receive notifications about tickets that are assigned to this topic by the users, or by the ticket handler team of EOSC-hub.

The Helpdesk therefore serves two groups, offering the following features to them:

- Main features offered to users:
  - Creation of a ticket for any of the EOSC Services (Hub and EOSC Portfolios)
  - Display all the tickets created by the owner
  - Find previously created tickets
  - Receive notifications about answers and changes to the tickets
  - Login with the EOSC AAI system
- Main features offered to the provider Helpdesk Team:
  - Notification when a new ticket is created
  - Classification of the tickets
  - Escalation of the tickets to the ticket handler team of EOSC-hub
  - Creation of a new support unit with assignation of an administrator role to specific users
  - Management of incident or disruption of Hub services
  - Interface for communicating with other service providers ticketing systems
  - First level support for EOSC integrated services as a service
  - Interface with a Known Errors Database and with a Change Management Database

EOSC services can use the EOSC Helpdesk choosing one of the following integration options:

1. Direct Usage: Use directly the EOSC helpdesk as the ticketing system for the service.
2. Ticket Redirection: Use the EOSC helpdesk only as a contact point to redirect the entry request for the specific service to a mailing list.
3. Full Integration: Integrate an external ticketing system with the EOSC helpdesk infrastructure to enable transfer of tickets between them.

3.4. Research data management

EOSC-hub offers various services for service providers to help them more easily manage research data and implement scenarios where research data (often ‘big data’) need to be stored, transferred, analysed, indexed, etc. Some of the typical situations when a service can benefit from such support:

- The service needs to store and analyse data with the use of external compute and storage resources (e.g. cloud resources)
- The service needs to copy/move big data between your premises and other sites
- The service needs an external repository to deposit research data or scientific applications for broader sharing and reuse

EOSC integration planEOSC integration plan

Figure 3-1. positions the relevant services along the 'virtuous cycle of research', showing the different services of EOSC-hub that can be used to

1. Discover existing research data and services/applications
2. Process and analyse digital research data
3. Archive and curate research data
4. Deposit and make available research data for reuse

These services, even if they are positioned within the same stage of the cycle, provide different features and/or are offered with different conditions of use.

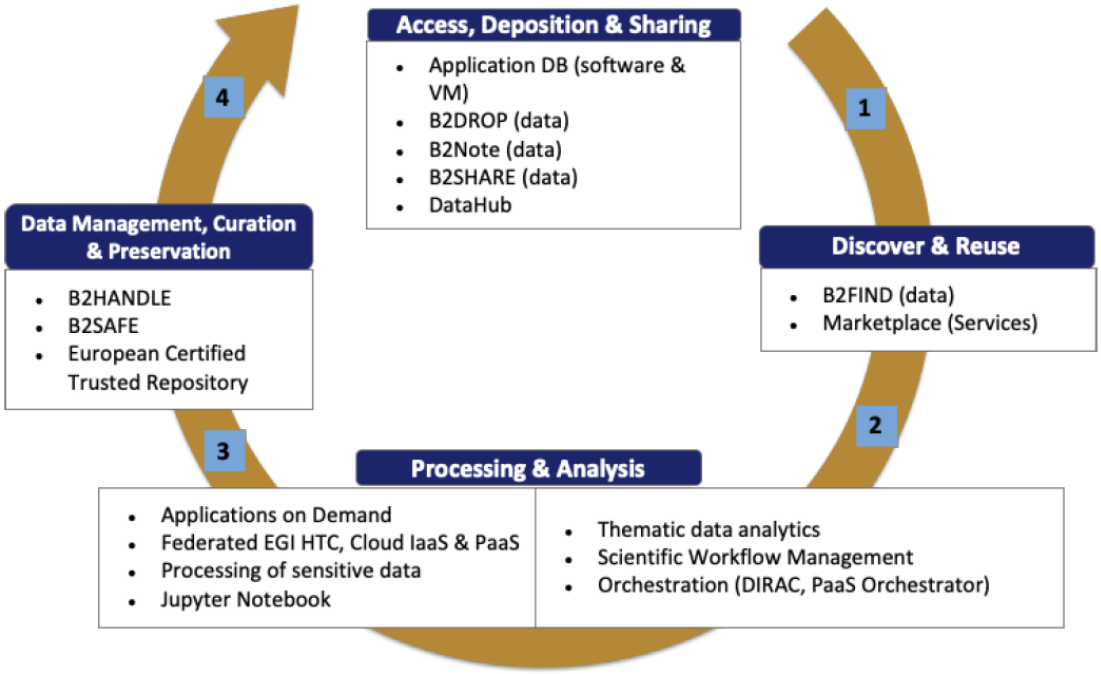

Figure 3-1. EOSC-hub services to support the virtuous cycle of research

The services can be all found in the EOSC Marketplace [14] and their Marketplace entries provide high level information, as well as pointers to user guides and other relevant materials.

3.5. Integrate with EOSC service management system

As part of the EOSC-hub contribution to EOSC, the project is developing and operating an IT Service Management system (SMS). The SMS ensures a robust and resilient service delivery in the EOSC federated infrastructure with different types of many-to-many relationships between users, providers, and clients.

What is IT Service Management?
The key idea behind IT service management could be summarized like this: By following a service-oriented approach, an IT organisation (which may be everything from an internal IT department over a shared IT unit up to an external IT provider) is able to better understand what they do and offer, and how this is aligned to the needs of their customers and users. IT service management (ITSM)

> refers to the entirety of activities – directed by policies, organized and structured in processes and supporting procedures – that are performed by an organization to design, plan, deliver, operate and control information technology (IT) services offered to customers. And by implementing IT service management processes, the activities carried out to plan, deliver, operate and control these services become more structured and repeatable, with clearly defined responsibilities. All this helps an IT organisation to increase their level of professionalism and organisational maturity.

The EOSC-hub SMS represents the entirety of activities performed by the providers that contribute to the EOSC core to plan, deliver, operate, and control the services offered to EOSC. The SMS also covers (to different extent) the activities of those service providers that have been onboarded to EOSC via the EOSC Portal. The activities carried out in the context of the SMS are structured and organised into processes and procedures according to the FitSM IT Management standard. By defining requirements, the 14 processes of FitSM help service providers:

Process	**Objective**
Service portfolio management (SPM)	To define and maintain a service portfolio
Service level management (SLM)	To maintain a service catalogue, and to define, agree and monitor service levels with customers by establishing meaningful service level agreements (SLAs) and supportive operational level agreements (OLAs) and underpinning agreements (UAs) with suppliers
Service reporting management (SRM)	To specify all service reports and ensure they are produced according to specifications in a timely manner to support decision-making
Service availability and continuity management (SACM)	To ensure sufficient service availability to meet agreed requirements and adequate service continuity
Capacity management (CAPM)	To ensure sufficient capacities are provided to meet agreed service capacity and performance requirements
Information security management (ISM)	To manage information security effectively through all activities performed to deliver and manage services, so that the confidentiality, integrity and accessibility of relevant information are preserved
Customer relationship management (CRM)	To establish and maintain a good relationship with customers receiving services
Supplier relationship management (SUPPM)	To establish and maintain a healthy relationship with suppliers supporting the service provider in delivering services to customers, and monitor their performance

Incident and service request management (ISRM)	To restore normal / agreed service operation within the agreed time after the occurrence of an incident, and to respond to user service requests
Problem management (PM)	To investigate the root causes of (recurring) incidents in order to avoid future recurrence of incidents by resolving the underlying cause, or to ensure workarounds/temporary fixes are available
Configuration management (CONFM)	To provide and maintain a logical model of all configuration items (CIs) and their relationships and dependencies
Change management (CHM)	To ensure changes to CIs are planned, approved, implemented and reviewed in a controlled manner to avoid adverse impact of changes to services or the customers receiving services
Release and deployment management (RDM)	To bundle changes of one or more CIs to releases, so that these changes can be tested and deployed to the live environment together
Continual service improvement management (CSI)	To identify, prioritize, plan, implement and review improvements to services and service management

For each of these processes, as well as for a number of general aspects in the context of ITSM, FitSM (within the FitSM-1 document) defines a small number of implementation requirements, while the FitSM-2 document provides guidelines on the activities to set up and implement ITSM using these processes. The FitSM-3 document describes the proposed roles to be assigned to execute the ITSM processes as part of a service management system.

At a base level, all onboarded services become in the scope of EOSC-hub Service Portfolio Management (SPM) when they are included into the EOSC Service Portfolio, and then publicly exposed in a Service Catalogue. How the scope of other EOSC-hub SMS processes impacts on new onboarded services depends on the choices the service providers make for integrating with other Hub Portfolio components. For example, enabling ‘ordering’ (i.e. users have to request access to the service via the EOSC Marketplace) will bring Your service partially into the scope of CRM, using the Helpdesk involves Your service in the ISRM process, and so on. Additional integration activities may bring the services within the scope of other SMS processes.

4. NEANIAS services to be integrated to EOSC

The following section introduces the NEANIAS services that will be integrated into EOSC. These services will cover three major resource sectors like Underwater, Atmospheric and Space research. Beyond these thematic services, NEANIAS develops/provides Core services in 4 groups. C1 covers Open-Science lifecycle support services, while C2 covers EOSC hub, RIs and cloud integration enabling services. The C3 AI and C4 Visualization services will support the NEANIAS thematic services as well, but the integration of C3 and C4 services to EOSC are optional and will be decided later. On the other hand, C1 and C2 provide basic services to enable integration of the thematic services in NEANIAS i.e. not services targeting reuse by 3rd parties and as such they are not presented in detail in this section. Nevertheless, some of those services have their own requirements and expectancies from EOSC and as such they are subject to suggestions presented in the previous section of the document.

4.1. Underwater services

Underwater surveys have numerous scientific and commercial applications in the fields of archaeology, geology, biology, and energy involving tasks such as ancient shipwreck prospection, ecological studies, environmental damage assessment, and detection of temporal changes. More specifically, this will be a significant easier entry point for the targeted users/customers going from raw data (U1) to maps, aided by an agile and fit-for-purpose service, allowing them to create higher resolution maps compared to DTMs shared by initiatives like EMODNET and GEBCO. Moreover, scientists and marine operations companies can benefit from acquired images as they provide, from the cognitive point of view (U2), the most precise and accurate representation of the areas surveyed, enabling a detailed analysis of the structures of interest. In addition, the novel seabed classification service (U3) will engage specific user-communities through the foreseen demonstrations in several underwater datasets addressing current needs of archaeologists, submarine volcano/geohazards, energy power cabling planners and oil & gas engineers.

4.1.1. U1 - Bathymetry Mapping from Acoustic Data

MB-System (Caress et al., 2018) is currently a popular desktop solution for post-processing bathymetry and backscatter from single- and multibeam echosounders. The package is programmed and constantly maintained under GNU GPL 3.0 (Open Source License) on Linux/Unix/Mac platforms and has several dependencies to other open source packages like GMT (Generic Mapping Tools), PROJ, OTPS (Tide models) and so on. With Jupyter Notebooks we see the opportunity to let users with different background handle bathymetric post-processing via predefined workflows. The cloud-based service created for this project would create an easy to use agile and fit-for-purpose workflow of multibeam data processing as simple as upload the data, choose and apply basic filters/corrections, define the desired resolution for grids/maps and retrieve the data products from the cloud server. Furthermore, the logging of activities and the potential to share Jupyter Notebooks could aid to cooperate e.g. between industry and research. Both software required to develop this service (MB-System and Jupyter Notebooks) are already existing and available since many years and meet the TRL 6 requirement, but further development time and resources are

required to piece them together to create the cloud-based service and validated with several datasets and application scenarios for the targeted TRL 8 level.

4.1.2. U2 - Seafloor Mosaicking from Optical Data

The current landscape of solutions for underwater optical mapping is still very tied to particular developments by a few research groups who create algorithms tailored to their particular acquisition hardware and target applications. As such, there is not yet any available attempt at dealing with a wide range of imaging conditions that are present in underwater surveys. The effort in service U2 is the first in its kind towards this goal powered by the cutting-edge infrastructures of EOSC. In particular, being the first of its kind, the impact of the service seems potentially significant. The key to success in this particular area is the ability to adapt to a wide range of conditions. We believe that the proposed approach, which is divided into an analysis and diagnostics task, followed by more specific processing steps, will be key to the success and adoption by a wide range of marine scientists and marine operations companies. Moreover, the target TRL is 8. The components required for this endeavor have been developed up to levels 5 (diagnostics task) and levels 6 (for the tasks of pre-processing and corrections, 2D mapping, 3D mapping). However, these have been demonstrated for particular applications and conditions, within the framework of their respective research projects. The challenge of the development behind this service is to make it robust to a wide range of input imaging conditions, which are representative of the intended science and commercial applications of the target user group.

4.1.3. U3 - Seabed Classification from Multispectral, Multibeam Data

U3 service focuses on the efficient exploitation of recent cutting-edge multibeam echosounders which go beyond the standard ones that collect backscatter at a single predefined frequency and can acquire backscatter at multiple spectral frequencies allows the acquisition of spatially and temporarily co-registered multispectral backscatter data offering capabilities for systematic and accurate mapping of the seabed. Based on recently developed classification framework which has been already validated in relevant environment (TRL 6) and in several datasets from the international 2018 R2Sonic Multispectral Challenge, the goal is to further optimize and develop an efficient and robust multispectral, multibeam data processing system integrating advanced machine learning tools for seabed classification. In contrast to similar R&D efforts in the proposed service different shallow and deep machine learning architectures will be further validated towards the classification of various seabed classes. An additional novelty lies in the optimization of a single classification model which will provide powerful pre-trained models for different sets of seabed nomenclatures. The validation process towards the targeted TR8 will be intensive against a number of datasets (e.g., USA, Canada, Atlantic coast, Mediterranean, Santorini, Kolumbo) demonstrating in a comprehensive way the applicability and scalability of the proposed cross-cutting service several application domains and user communities.

4.2. Atmosphere services

In the novel A1 service, although algorithms may be commercially available, no algorithms are standardized and widely accepted by the scientific and regulatory communities, raising a significant business opportunity since A1 allows currently missing regularization perspectives. The novel A2 service will engage geohazard scientific and government public protection authorities towards delivering timely alerts on forecasting geohazards activities.

For seismic/volcanic hazards, several exploitation opportunities will be raised, also from providing to stakeholders the crucial information of how gas and particulate concentrations change in space and time. The innovative A3 service will engage users and leverage business exploitation opportunities in (i) agriculture since air quality influences the agricultural productivity (ii) urban-related studies including planning, land use, and real-estate (iii) in insurance/ health since e.g. preventive measures in hospitals/ pharmacies can respond to citizens' needs.

4.2.1. A1 - Greenhouse Gases Flux Density Monitoring

The objective is to further optimize, automate and validate the existing software for flux density and fluxes estimations of gases, aerosol and energy from data obtained from specifically set meteorological stations, in order to integrate it as a novel cutting-edge service in EOSC. In particular, the front-end of the service will consist of a simple, user-friendly interface that will accept the data and guide the user in obtaining the desired flux densities, and energy balance results. It will be possible to upload the data by the user from his/her own database or from an EOSC database or from a linked to EOSC database. The user interface will efficiently demonstrate examples with several different challenging datasets which will target different user communities (national greenhouse gas emissions data validation, urban air quality authorities, meteorological services, and energy and power generating sector). The back-end of the service will consist of an orchestrator that will select the optimal method or combination of methods, for processing, depending on the available dataset. Two major micrometeorological pathways will be considered and integrated based on the aerodynamic gradient method and on the principle and application of the eddy covariance method. The examples will also demonstrate to the prospective users the attractiveness of the package and its ease of use. Finally, the validation towards TRL8 will include a large quantity of data and models from the community to evaluate and regulate the calculation of the above-mentioned flux densities and fluxes.

4.2.2. A2 - Monitor atmospheric Perturbations and Components in Active Tectonic Regions

The objective here is to further optimize and validate the existing data from specifically set of meteorological stations in order to determine possible correlations of gaseous and particulate components of the atmosphere with earthquake and volcanic processes. For seismic processes, the correlation will be established by quantifying the variation of a specific gas (i.e. radon) and particulate matter in the lowermost atmosphere before, during and after an earthquake event. This will offer a cutting-edge service to the international fundamental requirement and discussion on the importance and reliability of such gas/particulate variations as earthquake precursors phenomena. Regarding the volcanic processes, it is of paramount importance to quantify the input of magma-derived products in the atmosphere in the form of gases and ashes. In particular, the front-end of the service will consist of a simple, user-friendly interface that will accept the data and guide the user in obtaining the desired correlations of atmospheric components with the meteorological parameters. It will be possible to upload the data by the user from his/her own database or from an EOSC database or from a linked to EOSC database. A variety of real data/ cases will be employed for service validation progressing from a TRL6 to a TRL8 situation.

4.2.3. A3 - Air Quality Estimation, Monitoring and Forecasting

This new EOSC service will be delivered by further optimizing, automating and validating existing software modules for air quality estimation along with the novel integration of satellite, geospatial and in-situ data. The user interface will efficiently demonstrate examples with several different challenging datasets which will target different user communities (urban air quality authorities, meteorological services, energy and power generating sector, and industrial air pollutant emitters) The examples will also demonstrate to the user the attractiveness of the package and its ease of use and also how it can be used, not only, to improve urban planning, strategy and interaction with the city residents but also can be integrated with data from other urban environment areas (for instance, mobility) and provide different functions, such as alarms management, risks and emergency situations detection, geo-referencing data, statistical analysis and life quality reports. Furthermore, a Smart Air Quality's management platform as a SaaS platform with web responsive interfaces and REST interfaces for integration with external systems will be integrated. The dashboards on the platform will be configurable, being able to give daily, weekly and monthly reports of all the data collected by the sensing stations (even their battery status and management of the alerts regarding malfunctioning or failure). The validation will include several datasets, application cases progressing from TRL6 to TRL8.

4.3. Space services

The foreseen services will deliver a springboard of tools to enrich the workflows of a wide range of targeted users from academic/ research institutions to industrial stakeholders e.g. aerospace engineering and related technological companies, but also national space agencies and public outreach bodies e.g. space museums and planetariums. S1 will provide the required missing tools that enable the efficient and scalable visual discovery, exposed through advanced interaction paradigms exploiting virtual/ augmented reality. S2 will deliver multi-dimensional space maps through novel mosaicking techniques to a variety of prospective users/ customers (e.g., mining & robotic engineers, mobile telecommunications, space scientists). The delivered structure detection capabilities (S3) will leverage the targeted-users' opportunities for efficiently identifying and classifying specific structures of interest. All space services will provide a multitude of validated operational tools allowing a variety of space stakeholders to enhance their business portfolio.

4.3.1. S1 - FAIR Data Management and Visualization for Complex Data and Metadata

The Space environment shares different needs ranging from management of complex datasets, scientific visualization, visual analytics including combining data analytics and mining with visualization also exposing outputs within advanced interaction environments using virtual and augmented reality. The developed tools and services in Space will be tailored to 2D and 3D model computation, visualization and virtual reality navigation applications. VisIVO [22] and SPLOTCH [23] are suites of visualization software and services targeting the virtual observatory environment and exploiting HPC and distributed computing infrastructures. The Astronomical Data Navigator-ADN is a virtual reality application that allows realistic and precise visualization of astronomical catalogs in a virtual, navigable and interactive 3D environment. PlanetServer [24] allows planetary surface data access via client/server interactions and 2D/3D visualisation. These tools and services are all in TRL 6

status and will be ported to the EOSC, starting from experiences in previous projects (e.g. in the EOSCPilot project), by optimizing specific features as these are needed by the involved scientific groups. The S1 service will optimize data discovery under the FAIR principles by the publication of outputs through the IVOA framework, which is in very wide use by the astronomical community and is increasingly being adopted by the planetary community following the EuroPlanet, VESPA and PlanMap initiatives. Data accessibility will be guaranteed through deployment of standard protocols and interfaces implemented by the IVOA TAP and EPN-TAP protocols and the REST interface.

4.3.2. S2 - Map making and mosaicking for multidimensional image

Space science requires services for making high quality images from the raw data captured by instruments (map making) and for assembling those images into custom mosaics (mosaicing). There are some tools available for the community that are in TRL 6 and will be ported to the EOSC, optimizing the specific features as needed by the relevant scientific groups. Concerning radioastronomy data, interferometric observations, due to the peculiarities of data acquisition, the service will take into account primary beam effects as a weight for data to be mosaiced. Planetary mapping requires high image quality so that individual data products/granule are co-registered to achieve iteratively cartographic description (e.g. geologic) of a planetary surface. To achieve mapping there is need for higher-level products, that are made accessible (high-level data are not always available on public space agency archives, such as PDS/PSA) and co-registered, mosaicked as needed.

4.3.3. S3 - Structure detection on large map images with machine learning techniques

The currently experienced dramatic increases of Space data volumes make automatic structure detection a necessity. This need will become more and more important due to the continuous increase of data volumes that will need to be analysed. Recently project partners were involved in projects for automatic structure recognition such as FP7 ViaLactea project and SKA precursor activities. CAESAR [25] is a software tool for extraction and parameterization of both compact and extended sources present in astronomical maps. CuTEx [26] analyses images in the infrared bands and, in particular, it was designed to resolve problems concerning the study of star forming regions. Planetary exploration missions have the constant need for terrain characterization, largely based on orbital remote sensing data. S3 service will be focused on exploiting the existing TRL6 software to perform pattern and structure detection in astronomical surveys as well as in planetary surface composition, topography and morphometry. The service is expected to integrate cutting-edge machine learning algorithms, adopting pre-trained convolutional neural-networks for computer vision tasks (e.g. recognition, segmentation) adapted to the project-specific tasks by means of transfer learning approaches, to perform advanced classification for structures of sources in the sky or planetary surfaces to identify regions of interest.

4.4. C3 Artificial Intelligence services

C3 Artificial Intelligence services form the upper level of core services in the NEANIAS ecosystem and provide facilities that may reach up to the end user offering features of a typical Machine Learning (ML) workflow lifecycle, and that are also intended for composition of higher level services. In this subsection overall role of C3 services in NEANIAS is presented,

then each service is described in alignment with the Service Description template adopted by the project. The EOSC integration of the C3 services is optional and will be decided later.

4.4.1. C3.1 AI Science Gateway: service for development of ML models using Jupyter Hub

A development environment is necessary for the researchers to design, prototype, implement and test AI powered solutions. To serve as a universal core service for multiple users, the popular IPython based Jupyter Hub project is preferred, with TensorFlow and Keras libraries.

Researchers are able to prototype solutions, while computationally intense training can be loaded to a computational cluster. Results can be visualized using the Visualization gateway (C4.2).

4.4.2. C3.2 Serving trained ML models

Accessing trained models as web services should be possible seamlessly, with low efforts. To deploy a model as an API, large amount of memory is necessarily allocated and used constantly. Technically, there are multiple approaches to implement model serving, user and researcher-friendliness is prioritized.

The fundamental idea behind model serving is that authorized clients are able to access prediction functionality without directly accessing the model, which would create a serious overhead caused by data transfer and memory allocation. Instead, a RESTful API based communication is preferred between a model server and the client. The server implements the model, allocates memory to efficiently handle prediction, and listens for input parameters on the interface. On request, prediction is done using the inner implementation, and the results are responded to the client.

4.4.3. C3.3 Distributed Multi-GPU training of large ML models using Horovod

The training of deep neural networks on large amount of data requires a significant amount of processing time, which could be reduced by using GPU accelerators. The next step in increasing the performance is by using multiple GPU-enabled nodes in a distributed environment. While the parallel processing of deep neural networks has its limitations, Horovod [12] is a robust tool which can be used to scale the training of TensorFlow based network implementations.

4.4.4. C3.4 Distributed Machine Learning using SparkML

In addition to models based on neural networks (that can benefit from a multi-GPU deployment, especially for learning phases) that are managed through C3.3, other AI techniques for classification, regression, or clustering tasks can however benefit from a deployment and execution in a distributed and memory-based architecture such as the one offered by Spark. In particular, C3.4 will support a variety of machine learning models such as Support Vector Machines, Decision Trees and ensemble models (Random Forests and Gradient Boosting) for both classification and regression, clustering algorithms, and other support functions that are useful in real-world, large scale, machine learning pipelines. The set of basic ML algorithms and models can also be extended by means of creative parallel computing approaches (e.g. parallel implementations of DBSCAN algorithm on top of Spark, an algorithm which is not natively provided by MLib, are known and available).

4.5. C4 Visualisation services

C4 services, namely “visualisation services“, in the context of NEANIAS, form the upper level of core services in the NEANIAS ecosystem and provide facilities that may reach up to the end user offering features of a typical visualisation workflow lifecycle, and that are also intended for composition of higher level services.

In this subsection overall role of C4 services in NEANIAS is presented, then each service is described in alignment with the Service Description template adopted by the project. The EOSC integration of the C3 services is optional and will be decided later.

4.5.1. C4.1 Framework for Visual Discovery (VD)

C4.1 will provide tools to support data intensive computing for visual scientific discovery (including research training and outreach) for observational data, theoretical simulations and 2D/3D tiling and maps. C4.1 will consist of three distinct high-performance services for visual analytics from multidimensional data tables (VD-VisIVO), high-quality volume rendering of particle-based datasets exploiting a variety of parallel programming models leveraging upon heterogeneous infrastructures (VD-Splotch) and creation of interactive 2D/3D tilings and maps (VD-Maps). Details on these services are summarised in the tables below.

4.5.2. C4.2 Visualisation Gateway (VG)

C4.2 will provide a development environment for designing, rapid prototyping, implementing and fully testing complex visualisation solutions for realising common data exploration workflows. To serve as a universal core service for multiple users, the popular IPython based Jupyter Hub project has been selected. C4.2 will be built upon this and the framework for visual discovery developed in C4.1. C4.2 services are envisaged to be interconnected with C3.1 to visualise AI powered solutions, C4.3 to underpin powerful VR/AR solutions and C4.4. to facilitate end-user data accessibility. C4.2 are expected to be deployed in a way that is fully embedded within the relevant workflows of end-user activities, while exploiting diverse parallelisation models and infrastructure accelerator capabilities.

4.5.3. C4.3 Toolkit for Cross Realities (XR)

C4.3 will provide a toolkit underpinning an environment for designing, implementing and testing complex visualisation solutions exposed to end-users via Cross Realities (XR) mechanisms, e.g. those based on Virtual Reality (VR) and Augmented Reality (AR). To serve as a universal core service for multiple users, popular software frameworks and technologies have been selected based on a set of components used by an existing TLR6 software solution called Astro Data Navigator (ADN), e.g. the frameworks provided by the game engine Unity and HTC Vive headset. Such components will support interactive data exploration and navigation mechanisms e.g. for advanced comparisons in multidimensional and multifrequency datasets for research and public outreach. C4.3 services will be built upon extending components with a focus on providing services with enhanced realism, precision and usability along a number of aspects relating to enriched user experiences covering e.g. novel navigation mechanisms, advanced real-time tracking, VR/AR technologies and seamless integration of large-scale catalogues. C4.3 will exploit the framework for visual discovery developed in C4.1. C4.3 services are envisaged to be interconnected seamlessly with C4.2 to furnish complex visualisation workflows and C4.4 to underpin advanced data

access. The components identified so far are a tracking system interface integrating with motion tracking, a data connector for retrieving, in a generic way, data coming from different sources (individual files or databases) and a positioning manager providing the capabilities to retrieve specific data about position/rotation of particular objects for specific temporal instants. The implementation of the later is envisaged to deploy cloud infrastructures to extract the relevant information through Spice kernels. As part of the XR toolkit various VR/AR viewers are expected to be realised visualising not only space objects but also 3D data of various typologies of interest to other NEANIAS domains (e.g. atmospheric).

4.5.4. C4.4 Spatial Data Stores (DS)

Spatial data stores allow data to be referenced, query and retrieved, based on a given location (e.g. the globe, the Earth, planetary bodies with a proper reference system or the sky).

Earth Observation data, including Copernicus datasets and products, are referenced in the ADAM Data Access System (DAS), a software module that manages a large variety of geospatial information that feature different data format, geographic / geometric and time resolution. It allows accessing, visualizing, sub-setting, combining, processing, downloading all data sources simultaneously. The DAS exposes OGC Open Search and Web Coverage Service (WCS 2.x) interfaces that allow discovering available datasets and subset them in any dimension with a single query.

Planetary bodies differ from the sky on their concavity and dimensions, bodies being concave surfaces of limited sizes whereas the sky is convex and limitless. A coordinate reference system at the interface layer on spatial data bases is responsible to translate user requests in to internal data reference.

4.6. Identified requirements/interfaces

Based on the investigation of the services listed above, we identified the following services/support as requirements to integrate the thematic services to EOSC:

- Service modelling
- Authentication and authorization
- Monitoring
- Accounting
- Helpdesk
- Data Management Planning
- Data Catalogue

In the next section, we intend to introduce the available EOSC services in these categories.

5. Overview of existing EOSC services

In this section, already existing EOSC services will be introduced which could be used and integrated with the NEANIAS services. These services have been selected based on integration requirements of EOSC, and based on the specific needs of the NEANIAS thematic services. We aim to onboard the NEANIAS thematic services in EOSC (i.e. publishing them in the EOSC Portal), and we wish to expand their capabilities with the use of the EOSC-hub federation services, and some of the Research Data Management services. The next subsections provide an overview of these specific services.

5.1. EOSC Portal

NEANIAS aims to deliver thematic services to researchers via EOSC, therefore the registration of these services in the EOSC Portal is a compulsory step. The EOSC portal is the European Open Science Cloud gateway offering researchers the ability to discover, access and compare EOSC services and resources. The EOSC portal provides access to all researchers from all domains and is open to scientific users and service providers.

The EOSC portal brings together services and resources such as computing, storage, data management, networking, research publications and software from national and international research infrastructures, organizations and more. The EOSC portal Catalogue and Marketplace give researchers the ability to:

- Browse, search and discover a variety of services and resources
- Compare different services
- Access the services selected easily with existing credentials from a very broad set of academic, community or social identity providers
- Find out detailed information on each service including access policies, maturity level, latest updates, country availability, languages supported and more.
- Receive help and support for all services
- Write reviews for all services

Providers of services are able to contribute their services and resources to the EOSC portal following the onboarding process. Currently, the process for service onboarding [27] is providers submit a request for onboarding via a dedicated form, then fill in the Provider and Resource profiles templates with information on their profile and services and a validation team reviews the information submitted to ensure the rules of participation are met. The service is entered in the Service Portfolio and a subset of data is transferred to the Marketplace in EOSC portal.

This process is continuously evolving and new sets of rules and criteria are being examined and introduced by the EOSC-Hub and EOSC-Enhance projects.

5.2. Federation Services

In this subsection, already existing EOSC federation services will be introduced, which we plan to integrate with the NEANIAS services.

5.2.1. AAI services

The NEANIAS thematic services require user login and authorization. The project is working on a NEANIAS AAI system that will simplify the user AAI and that can act as a connection point between the NEANIAS thematic services and the EOSC AAI. The EOSC AAI system acts as a bridge between thematic services and a broad set of identity providers, enabling users to login to the thematic services from consist of a portfolio of endpoints. Establishment of 1-1 connection between the service provider and NEANIAS AAI team shall be made. The NEANIAS AAI team should choose one of these endpoints and connect to it:

- EGI Check-in [15] is a proxy service that operates as a central hub to connect federated Identity Providers (IdPs) with EGI service providers. Check-in allows users to select their preferred IdP so that they can access and use EGI services in a uniform and easy way. Main characteristics:
    - Enables multiple federated authentication sources using different technologies
    - Increased productivity and security
    - Federated in eduGAIN as a service provider, publishing REFEDS RnS and Sirtfi compliance
    - User registration portal to allow accounts-linking
    - Combines user attributes originating from various authoritative sources (IdPs and attribute provider services) and delivers them to the connected EGI service providers in a transparent way.
- The B2ACCESS AAI proxy service [16] is arbitrating access to registered Service Providers via different protocols. Integrated Service Providers consume Attribute assertions from the B2ACCESS service when the End User accesses them. B2ACCESS allows EUDAT users to authenticate themselves using a variety of other Identity Providers or credentials. The features of B2ACCESS:
    - supports several methods of authentication via the users' primary identity providers (OpenID, SAML, x.509)
    - can be used as primary identity provider, if necessary
    - can be integrated with any service of the CDI service provider federation
    - is integrated with EduGain and supports identities from theoretically hundreds of Universities and Research institutions worldwide.
    - provides unique and persistent EUDAT-wide meaningful identifiers.
- The eduTEAMS [29] service enables research communities to securely access and share common resources and services. Leveraging the ubiquitous presence of eduGAIN federated identities, eduTEAMS enables communities to securely authenticate and identify their users, organize them in groups, assign them roles and centrally manage access rights for using community resources. As research is not confined only in the research institutes and universities, eduTEAMS caters also for users coming from the industry or citizen scientists who may not have access to eduGAIN. It does so by supporting external (non-eduGAIN) identity providers, such

as social networks providing federated identities, community identity providers and other platforms that can provided federated users identities.

- The INDIGO Identity and Access Management Service (INDIGO-IAM) [30] provides a layer where identities, enrolment, group membership and other attributes and authorization policies on distributed resources can be managed in an homogeneous way, supporting identity federations and other authentication mechanisms (X.509 certificates and social logins). The IAM service has been successfully integrated with many off-the-shelf components like Openstack, Kubernetes, Atlassian JIRA and Confluence, Grafana and with key Grid computing middleware services (FTS, dCache, StoRM).

5.2.2. Availability monitoring

Availability monitoring is a service where users and service providers can check the status of an EOSC service and get information about its availability and reliability in the selected time period. In EOSC the following monitoring services can be found which are potential to be utilised for NEANIAS monitoring.

- EGI Service Monitoring [17] is an infrastructure monitoring service which relies on data generated by functional probes. The features of the EGI Service Monitoring:
  - Minimal development effort for setting up monitoring services
  - Read-to-use user interface
  - Automated reporting tools
- PerfSONAR [18] is a network monitoring service aimed at the research and education community. It is based on an open-source, modular and flexible architecture, which accounts for the complexity of network infrastructures. The features of the PerfSONAR service:
  - a dashboard which reports matrixes of worldwide network monitoring measurements
  - uniform interfaces for data retrieval and uniform measurements formats
  - user-defined scheduled measurements
  - on-demand measurements
- Although ARGO monitoring [19] at the time of writing is not available as a federated EOSC service, it is currently employed for the monitoring of the EOSC Portal. The ARGO monitoring and alerting service provides the following features:
  - Open source modular architecture
  - Horizontal scalability
  - Scalable elaboration of metrics

NEANIAS will provide an availability monitoring instrument in order to be able to verify its compliance to the intended Services SLA. Integration with a monitoring infrastructure from EOSC will be reconsidered at later stages of the project, taking into account the evolution of the relevant policy and tools in EOSC.

5.2.3. Accounting

The NEANIAS thematic services will consume storage and compute capacity from underlying provider(s). Having a full understanding of this use (e.g. on a per provider and per user basis)

would help the project build the picture of its operational costs and plan sustainable business model for beyond the project lifetime. The EOSC Accounting system can be helpful here and may offer the required level of details about our compute and storage use. The EOSC Accounting system is based on the EGI Accounting system, called ARGO:

- EGI Accounting [10] stores user accounting records from various services, such as Cloud, HPC and storage usage. It works thanks to a network of message brokers that transfer usage data from the host to a central repository of information. The data is handled securely and can be consulted online through the EGI Accounting Portal [11]. EGI Accounting provides:
    - Increased control over resource consumption
    - Reduced overhead of defining data models, architecture, and setup of an accounting system
    - Reduced cost of maintaining an accounting infrastructure
    - Access to a reliable, high available, high performance service
    - User friendly web interface

However, accounting spans largely other models rather than direct measurement of physical (or virtualized) resources and NEANIAS is in the process of defining a model for those too, as it is facing EOSC landscape from the perspective of complete financial sustainability. As verified also via latest EOSC calls, utilization of resources is only one aspect of many.

Regarding the Accounting Policy, there are two challenges:

- NEANIAS employees a service rather than an infrastructure provided perspective, which needs additional accounting concepts (e.g. rated API calls, logons, publication sets etc)
- NEANIAS engages a number of infrastructure providers that employ substantially different infrastructures (that may be further abstractions of other lower level resources), technologies and policies. The integration of proxies may be neither feasible nor acceptable due to policies of those providers.

Furthermore, NEANIAS by definition is not specifically targeting accounting integration and its accounting policy is targeting project-specific KPIs. As such the project shall proceed with the internal accounting approach and will further analyse the feasibility of resource provider level accounting integration at later stages, closely monitoring EOSC policies, concepts and services evolution in the area.

5.2.4. Helpdesk

The EOSC Helpdesk system is a ticketing service that connects EOSC users and service providers. NEANIAS could leverage this system to respond to user requests and other types of trouble tickets arising the EOSC context. Helpdesk services are not generic and reusable, but instead aimed at providing support for specific EOSC services.

- EGI Helpdesk [20] provides you with the information and support you need to troubleshoot your product and service problems. You can report incidents, bugs or change requests. EGI provides support to users and operators through a distributed helpdesk with central coordinating. The central helpdesk provides a single interface

for support. The support activities are grouped in first and second level support. Main characteristics are:

- Central point of contact for support
- Repository of information and solutions
- Keeps track of progress of ongoing issues happening on the infrastructure

- Open Science Helpdesk [21] provides support for Open Science issues to European Research Institutions, projects and institutions. It has the following features:
  - Runs 24x7
  - Supported by FAQs, Factsheets, Briefing Papers and topical webinars
  - Ran by National Open Access Desks (NOADs) at national level

NEANIAS has no specific mandate for its service providers with respect to ticketing, so each one of them may choose the ticketing option they wish to adopt. The requirement set by the project is that:

- Access is open to all users
- SSO should be adopted.
- Metrics should be shared with the project in order to assess service quality.

Apart from the aforementioned options, NEANIAS is offering a ticketing system to all its services that covers those requirements.

5.3. Research Data Management

In this section we provide an overview of some of the existing EOSC services regarding FAIR data management that we opted to reuse and integrate, so that NEANIAS services are conformant with the Open Data Guidelines [2] and the FAIR principles [8]. As our project moves along and as the EOSC service portfolio expands we'll also expand our view and will consider additional research data management services for integration.

5.3.1. ARGOS Data Management Plan tool

Argos [3] is an open and extensible EOSC service that simplifies the management, validation, monitoring and maintenance and of Data Management Plans. It is a service provided by OpenAIRE, which very recently entered its public access stage. Argos allows actors (researchers, managers, supervisors, etc.) to create actionable DMPs that may be freely exchanged among infrastructures for carrying out specific aspects of the Data management process in accordance with the intentions and commitment of Data owners. To facilitate the adoption of the platform, NEANIAS has engaged OpenAIRE experts since the beginning of the project to provide support on the use of the platform both in live presentations and as web training.

Argos is based on OpenDMP [5] which is by itself a free and open source software used for actionable machine Data Management Planning, that closely follows RDA DMP WG developments and specifications [4].

In NEANIAS, we exclusively use Argos to maintain our Data Management plans.

5.3.2. Research Data Catalogue

Zenodo [6], an EOSC service, is the catch-all repository developed and maintained by CERN as part of the OpenAIRE project, in support of Open Science and FAIR principles. Zenodo's aim is to store all of its contents for the lifetime of the repository, providing free access to all metadata under the CC0 license (except for email addresses). Additionally, all metadata can be harvested via the OAI-PMH mechanism for repository interoperability [7]. Moreover, Zenodo offers a REST API [9] to allow applications to:

a) upload and publish research outputs
b) search published records
c) upload and download files

In NEANIAS, we opted to use Zenodo via its REST API for all research product metadata, and for all open-access research data. Additionally, we will use Zenodo as a Persistent Identifier (PID) production service, to allow NEANIAS services to automatically obtain PIDs for research outputs.

6. Integration plan to EOSC

In this section we present the integration plan of NEANIAS services to EOSC. Our fundamental goals are:

1. Publish NEANIAS thematic services in EOSC. Consider the publishing of additional NEANIAS services in EOSC in a future stage.
2. Integrate NEANIAS services (thematic and core) with relevant EOSC services where we see value of the integration. (e.g. to enhance our operations, to deliver more features to users)
3. Develop the NEANIAS IT Service Management System gradually, and in an EOSC compatible way.

We aim to achieve these goals following an agile approach that is based on successive development cycles, considering the following:

- The cycles bring NEANIAS services from very basic EOSC integration with minimal functionalities to more complex EOSC integration with advanced capabilities.
- We specify the exact content for each cycle considering the situation at that moment when we reach that cycle. Therefore: Cycle 1 (which happens first) should be clear in this deliverable, while cycle 2 can be fuzzier, because we can revisit this once we see movement and progress in Cycle 1. Cycle 3 can be even fuzzier than cycle 2 and will be clarified once Cycle 2 is finished.
- The cycles can be repeated for the different Thematic Services independently from each other. For example, cycle 1 can start for the most mature Thematic service as soon as that thematic service is ready for publication in EOSC. Other thematic services will be considered for cycle 1 once they become mature enough for EOSC.
- The project should start cycle 1 as soon as possible for the most mature Thematic service. This will bring useful experience to the project concerning the EOSC integration work, which can be fed back to the other Thematic services.
- Cycle 1 as well as cycle 2 has a set of preparatory steps that NEANIAS should complete before any of the thematic services reach those cycles. (These preparatory steps should be completed only one time, while the cycles should be completed one time for each thematic service.)

Preparatory steps for cycle 1 (done once):

- NEANIAS portal and services catalogue prototype deployment
- NEANIAS service catalogue sign up and registration
    - Providers registration based on EOSC onboarding process
    - NEANIAS services terms of use definition and acceptance based on EOSC RoP
- Setup the NEANIAS IT Service Management skeleton where service management documentations can be developed and stored (to be defined by WP7).

Provisional timeline: Started at M3; Finish by M24;

EOSC integration planEOSC integration plan

Cycle 1 – Make NEANIAS thematic services accessible to EOSC users by following the EOSC onboarding process (done for each thematic service):

- Ensure EOSC RoP compliance (i.e. we are able to fill the EOSC Service Description Template for the NEANIAS thematic service which is in scope)
- Define NEANIAS services pricing as well as respective SLA(s) and OLA(s) for the service
- Activate ordering (access requests) from the EOSC Portal to NEANIAS thematic services
- Ensure EOSC access policies and licensing compliance
- Enable single sign-on from the EOSC Portal to the NEANIAS Thematic service (through the NEANIAS AAI service)
- Provide minimum level documentation and helpdesk support for users
- Ensure minimum level of user feedback collection and continuous improvement (CRM)

Provisional timeline: Start at M12; Finish by M31 (EOSC catalogue integration and AAI integration by M25 based on D6.1);

Preparatory steps for cycle 2 (done once):

- Release the NEANIAS research product catalogue based on OpenAIRE's Zenodo platform
- Release NEANIAS PID service based on OpenAIRE's Zenodo platform
- Release the NEANIAS Data Validation Service
- Release the NEANIAS Data Publishing Service
- Integrate with OpenDMP deployed on OpenAIRE as Argos

Cycle 2 – Enrich NEANIAS services with research product catalogue and data management capabilities (done for each thematic service):

- Integrate with the NEANIAS research product catalogue based on OpenAIRE's Zenodo platform
- Integrate with NEANIAS PID service based on OpenAIRE's Zenodo platform
- Integrate with the NEANIAS Data Validation Service
- Integrate with the NEANIAS Data Publishing Service
- Integrate with OpenDMP deployed on OpenAIRE as Argos

Provisional timeline: Finish by M32;

Cycle 3 – Improve support for EOSC users based on initial experiences (done for each thematic service):

- Provide FAQs about the services
- Activate NEANIAS usage accounting
- Activate NEANIAS monitoring
- Integrate NEANIAS Monitoring with EOSC Availability and Reliability monitoring
- Integrate NEANIAS usage accounting with EOSC Accounting

EOSC integration planEOSC integration plan

- Define NEANIAS services CRM processes
- Add support for customised SLAs and OLAs
- Improve software documentation
- Extend NEANIAS Helpdesk

Provisional timeline: Finish by M32;

The EOSC-hub integration handbook provides 4 groups of services that NEANIAS can integrate with. We aligned the NEANIAS services with these 4 groups, indicating which NEANIAS service should be interfaces/connected to which EOSC equivalent:

EOSC service areas		**NEANIAS Service**
1. EOSC Portal		Thematic services
2. Federation services	EOSC AAI Availability and reliability monitoring Usage accounting (compute and storage) Helpdesk (ticketing system)	NEANIAS AAI NEANIAS specific test probes Accounting Helpdesk support staff
3. Research data management services	Compute, storage, transfer, analysis, PIDs, etc.): Zenodo (research output catalogue + PID)	GARR Cloud Service Catalogue Research product Catalogue Data Validation
4. Service Management System (SMS)	14 processes and respective implementation tools	NEANIAS SMS (Processes for Order management, SLAs, OLAs, CRM…)

7. Conclusions

The first version of EOSC was opened in 2018, but the system is still in rapid development. EOSC-hub and OpenAIRE-Advance, in collaboration with other EOSC projects, put in place the EOSC Portal, onboarding process, federation services and a research data management services that NEANIAS can rely on. Based on the currently available EOSC capabilities in this deliverable we defined our intended use of these:

1. We will onboard the NEANIAS thematic services in the EOSC Portal. As next step we need to start populating the EOSC Service Description Template (SDT) with information from the thematic providers. We will do this by identifying the most advanced thematic service and will use it to test-run the EOSC SDT process.
2. We will interface NEANIAS core services with the EOSC federation services. We need to establish 1-to-1 connections between the respective EOSC and NEANIAS teams. (i.e. AAI, Monitoring, Accounting, Helpdesk)
3. We need to develop the initial version of some of the IT service management processes to be ready for handling orders, trouble tickets, user feedback. For this we need to put in place the NEANIAS SMS (Service Management System) skeleton.
4. We will integrate NEANIAS services with some of the already available Research Data Management service of EOSC. We have to further analyse the exact demands for these, and we will have to establish the connections between the respective NEANIAS service developer and EOSC service provider teams to realise the integrations. The work on data management integration is in progress and the policies will be presented internally after the 1st full release and will be presented in the next version of the deliverable.

H2020 projects, Horizon Europe projects and a growing number of national initiatives are expected to evolve the EOSC landscape in the coming years. NEANIAS will therefore stay open for new integration opportunities offered by EOSC, and reconsider/update its own integration roadmap as needed. D8.3 (due at month 22) will be the next formal checkpoint for NEANIAS to update the integration roadmap.

References

[1] EOSC-hub D4.1 Operational requirements for the services in the catalogue
[2] EC Digital Future Policies: Open Data. [https://ec.europa.eu/digital-single-market/en/open-data]
[3] ARGOS service [https://argos.openaire.eu/]
[4] Open Research Data Pilot in H2020 [https://www.openaire.eu/ordp/ordp/pilot]
[5] OpenDMP [https://gitlab.eudat.eu/dmp/OpenAIRE-EUDAT-DMP-service-pilot/]
[6] Zenodo [https://zenodo.org/]
[7] OAI-PMH [https://www.openarchives.org/pmh/]
[8] Wilkinson, M. D. et al, (2016). The FAIR Guiding Principles for scientific data management and stewardship. *Sci Data* 3**,** 160018 (2016). [DOI: https://doi.org/10.1038/sdata.2016.18]
[9] Zenodo REST API [https://developers.zenodo.org]
[10] EGI Accounting [https://eosc-hub.eu/support-services/Accounting]
[11] EGI Accounting portal [https://accounting.egi.eu/]
[12] Horovod [https://eng.uber.com/horovod/]
[13] EOSC-hub Helpdesk [https://helpdesk.eosc-hub.eu/]
[14] EOSC Marketplace [https://marketplace.eosc-portal.eu/]
[15] EGI Check-in [https://www.eosc-hub.eu/services/EGI%20Check-in]
[16] B2ACCESS AAI proxy service [https://marketplace.eosc-portal.eu/services/b2access]
[17] EGI Service Monitoring [https://marketplace.eosc-portal.eu/service/egi.egi_service_monitoring]
[18] PerfSONAR [https://marketplace.eosc-portal.eu/services/perfsonar]
[19] ARGO monitoring [https://argoeu.github.io/]
[20] EGI Helpdesk [https://marketplace.eosc-portal.eu/services/egi-helpdesk]
[21] Open Science Helpdesk [https://marketplace.eosc-portal.eu/services/open-science-helpdesk-c5811d6d-4d57-4e0a-8e1b-485db4a9ac12]
[22] E. Sciacca et al., "VisIVO Workflow-Oriented Science Gateway for Astrophysical Visualization," 2013 21st Euromicro International Conference on Parallel, Distributed, and Network-Based Processing, Belfast, 2013, pp. 164-171, doi: 10.1109/PDP.2013.31.
[23] M. Rivi, C. Gheller, T. Dykes, M. Krokos, K. Dolag: GPU accelerated particle visualization with Splotch, Astronomy and Computing, Volume 5, 2014, Pages 9-18, ISSN 2213-1337, https://doi.org/10.1016/j.ascom.2014.03.001.
[24] R. Marco Figuera, B. Pham Huu, A.P. Rossi, M. Minin, J. Flahaut, A. Halder: Online characterization of planetary surfaces: PlanetServer, an open-source analysis and visualization tool, Planetary and Space Science, Volume 150, 2018, Pages 141-156, ISSN 0032-0633, https://doi.org/10.1016/j.pss.2017.09.007.
[25] S. Riggi, A. Ingallinera, P. Leto, F. Cavallaro, F. Bufano, F. Schillirò, C. Trigilio, G. Umana, C. S. Buemi, R. P. Norris, Automated detection of extended sources in radio maps: progress from the SCORPIO survey, Monthly Notices of the Royal Astronomical Society, Volume 460, Issue 2, 01 August 2016, Pages 1486–1499, https://doi.org/10.1093/mnras/stw982

[26] Molinari, S., Schisano, E., Faustini, F., Pestalozzi, M., di Giorgio, A. M., & Liu, S. (2017). CUTEX: CUrvature Thresholding EXtractor. ascl, ascl-1708.
[27] EOSC-hub Service Provider Onboarding [https://wiki.eosc-hub.eu/display/EOSC/Service+Provider+Onboarding+-+overview]
[28] EOSC-hub Authentication and Authorisation Infrastructure, [https://confluence.egi.eu/display/EOSC/Authentication+and+Authorization+Infrastructure+-+AAI]
[29] eduTEAMS [https://wiki.geant.org/display/eduTEAMS/eduTEAMS+Home]
[30] INDIGO-IAM [https://indigo-iam.github.io/docs/v/current/about.html]