H2020-INFRAEOSC-2018-3

NI4OS-Europe

National Initiatives for Open Science in Europe

Deliverable D5.2
First report on provider and repository integration

Lead beneficiary(s):	IPB
Author(s):	NI4OS-Europe consortium
Status –Version:	Final version – g
Date:	January 28, 2021
Dissemination Level:	Public

Abstract: Deliverable D5.2 - First report on provider and repository integration - presents the current resource and provider on-boarding status. The deliverable gives details on the on-boarding procedure and reports fully and partially integrated repositories, thematic and generic services, as well as services from the pre-production environment.

The NI4OS-Europe Consortium consists of:

GRNET SA	Coordinator	Greece
ATHENA RC	Beneficiary	Greece
CYI	Beneficiary	Cyprus
UCY	Beneficiary	Cyprus
IICT	Beneficiary	Bulgaria
SRCE	Beneficiary	Croatia
RBI	Beneficiary	Croatia
KIFU	Beneficiary	Hungary
DE	Beneficiary	Hungary
ICI BUCURESTI	Beneficiary	Romania
UEFISCDI	Beneficiary	Romania
ARNES	Beneficiary	Slovenia
UMUKM	Beneficiary	Slovenia
IPB	Beneficiary	Serbia
UOB	Beneficiary	Serbia
RASH	Beneficiary	Albania
UNI BL	Beneficiary	Bosnia-Herzegovina
UKIM	Beneficiary	North Macedonia
UOM	Beneficiary	Montenegro
RENAM	Beneficiary	Moldova (Republic of)
IIAP NAS RA	Beneficiary	Armenia
GRENA	Beneficiary	Georgia

Document Revision History

Date	Issue	Author/Editor/Contributor	Summary of main changes
December 1, 2020	a	Dusan Vudragovic	Initial ToC
December 7, 2020	b	Dusan Vudragovic	Final ToC
December 21, 2020	c	NI4OS-Europe consortium	Initial input provided by the partners
January 15, 2021	d	Brina Klemencic, Andreas Athenodorou, Petar Jovanovic, Anastas Mishev, Dusan Vudragovic	Additional input provided by the partners and WP5 task leaders
January 27, 2021	e	Ognjen Prnjat, Anastas Mishev, Kostas Koumantaros, Dusan Vudragovic	Final draft
January 28, 2021	f	Milutin Radonjic	Internal quality check
January 28, 2021	g	Dusan Vudragovic	Final version

Table of contents

References

[1] Project NI4OS-Europe-857645 - Annex I - Description of the Action.

[2] Best practices for on-boarding and related policies, NI4OS-Europe deliverable D3.1, DOI: http://doi.org/10.5281/zenodo.3736143.

[3] Provider landscape analysis and provider categorization, NI4OS-Europe deliverable D5.1, DOI: https://doi.org/10.5281/zenodo.3932941.

[4] User communities' test-cases, NI4OS-Europe deliverable D6.3, DOI: https://doi.org/10.5281/zenodo.3971647.

[5] First report on pre-production environment, NI4OS-Europe deliverable D3.2, DOI: https://doi.org/10.5281/zenodo.3932925.

[6] EOSC Profiles, version 3.00, https://eosc-portal.eu/providers-documentation/eosc-provider-portal-resource-profile.

[7] Agora service management tool web page, https://grnet.github.io/agora-sp/.

[8] Directory of open access repositories - OpenDOAR, https://v2.sherpa.ac.uk/opendoar/.

[9] European open science infrastructure for open scholarly and scientific communication - OpenAIRE, https://www.openaire.eu/.

[10] NI4OS-Europe Agora user interface, https://agora.ni4os.eu/.

[11] NI4OS-Europe on-boarding dashboard, https://catalogue.ni4os.eu/.

[12] NI4OS-Europe helpdesk system, https://helpdesk.ni4os.eu/otrs/.

[13] OTRS helpdesk system web page, https://otrs.com/.

[14] NI4OS-Europe configuration management database, https://gocdb.ni4os.eu/portal/.

[15] NI4OS-Europe Agro monitoring service, production instance, https://argo.ni4os.eu/.

[16] NI4OS-Europe Agro monitoring service, development instance, https://argo-devel.ni4os.eu/.

[17] Monitoring guide for resource providers, https://wiki.ni4os.eu/index.php/Monitoring_guide_for_SPs.

[18] NI4OS-Europe Argo poem library, https://poem.ni4os.eu/ui/public_home.

[19] NI4OS-Europe accounting system, https://accounting.ni4os.eu/.

List of Figures

List of Tables

List of Acronyms

AAI	Authentication and Authorization Infrastructure
API	Application Program Interface
CPU	Central Processing Unit
EOSC	European Open Science Cloud
FitSM	Federated IT Service Management
GOCDB	Grid Operations Configuration Management Database
GPU	Graphics Processing Unit
HPC	High Performance Computing
HTC	High Throughput Computing
JSON	JavaScript Object Notation
OTRS	Open Ticket Request System
REST	Representational State Transfer
SLURM	Simple Linux Utility for Resource Management

Executive summary

What is the focus of this Deliverable?

The main focus of this deliverable [1] is to present practical details on the resource and provider on-boarding procedure, as well as to report the current status of the on-boarding activities. The status is presented per each resource category, for repositories, thematic and generic services, and services from the pre-production environment by listing fully and partially integrated resources, as well as providers.

What is next in the process to deliver the NI4OS-Europe results?

Deliverable D3.1 - Best practices for on-boarding and related policies [2] - sketched up initial on-boarding procedure in order to identify a set of services and tools relevant for the setup of an environment (pre-production environment) that will enable on-boarding activities. More details and methodology were presented in the deliverable D5.1 - Provider landscape analysis and provider categorization [3] - where different on-boarding aspects, requirements, and categorizations of available resources were analyzed. The output of such analysis of different aspects of resource description was the initial on-boarding plan, which did not include the scientific impact of a particular resource. The scientific impact of the particular service was assessed within the WP6 and reported in the deliverable D6.3 - User communities' test-cases [4]. Merging these aspects, more technical from the D5.1, and more scientific from D6.3, resource candidates were selected and on-boarded as it is presented in this deliverable. Therefore, this document reflects the current consortium's on-boarding experience, which, since it develops over time, will be revised in the upcoming deliverables D5.3 - Update of service catalogue, and D5.4 - Second report on provider and repository integration and horizontal service delivery. The content of this deliverable is relevant to the whole consortium. In particular, to WP3 and WP4 in the adaptation of existing and the development of new tools and policies for federation and on-boarding, which will be reported in the deliverable D3.4 - Best practices for on-boarding and related policies 2nd version, as well as to WP6 in the involvement and support of real user communities in testing and fine-tuning of the proposed solutions, which will be reported in the deliverable D6.4 - Service evaluation by user communities.

What are the deliverable contents?

The report on the resource and provider on-boarding status is organized as follows:

- Section 2 explains practical details of the established on-boarding procedure. It describes relations between the project's catalogue, its frontend implemented as the on-boarding dashboard, and the central EOSC catalogue. Within this section, we present the usage of the services from the pre-production environment and explain how resources are integrated with them. What is considered as the on-boarded resource in the current project stage is discussed in this section as well.
- Section 3 presents the current status of the provider on-boarding. This section lists mandatory information to be supplied by providers in order to be marked as on-boarded, and notes on-boarded providers together with the links to a full description at the project's on-boarding dashboard portal.

- • Section 4 gives details on the mandatory information expected to be supplied in order to consider a resource as on-boarded. Plans for further development are also described in this section, while the details about on-boarded and candidate resources are organized per resource type: repositories, thematic and generic services, and services from the pre-production environment.

The overall status of the on-boarding and related overall statistics is provided in Section 1, while Section 5 presents the deliverable conclusions.

Conclusions and recommendations

According to the initial on-boarding plan presented in the deliverable D5.1 - Provider landscape analysis and provider categorization [3], at the beginning of the year 2021, it was expected to have fully on-boarded 2 repositories, 2 thematic, and 2 generic services to the EOSC. We managed to accomplish this by on-boarding 2 repositories, 4 thematic, and 4 generic services. The initial plan also defined that the total number of resources registered within the catalogue will be 6 repositories, 12 thematic, and 12 generic services, which was fully satisfied as well by registering 7 repositories, 12 thematic, and 12 generic services. In addition, all services from the pre-production environment (7 services in total) were also on-boarded, so the total number of resources available within the catalogue is currently 38. During the on-boarding process, we have identified several practical issues and recommendations which are discussed in the deliverable, and which will be addressed in the upcoming period.

1. Introduction

An initial on-boarding procedure was set up in the deliverable D3.1 - Best practices for on-boarding and related policies [2] - where we identified the main steps in the process of on-boarding resources and providers, a request for on-boarding, information gathering, integration, validation, and publication, as well as tools to support each step in such process. These supporting tools are unified within the project's pre-production environment, which was reported in the deliverable D3.2 - First report on pre-production environment [5].

The initial on-boarding procedure was further developed in the deliverable D5.1 - Provider landscape analysis and provider categorization [3] - where it was equipped with the on-boarding methodology that allows us to perform categorization of available resources in the region. After the landscape analysis, we have collected 97 resource descriptions out of which 27 are generic services, 38 are thematic services, and 32 are repositories. These numbers increase with time and regularly new candidate resources (mainly thematic services) emerging from the consortium are added. Therefore, our pool of resource candidates for the on-boarding grows over time, while we aim to select the resources with the highest relevance and impact to be integrated into the EOSC. Hence, the identification of resources that have an added value for the community and users, and that address needs of a potentially large customer base, is one of our main priorities, which drives the selection of 20 generic services, 20 thematic services, and 15 repositories (KPIs defined in the DoA) to be on-boarded into the EOSC within the framework of the NI4OS-Europe project. Based on the EOSC profiles specification [6], we created areas within which an assessment of a particular resource could be performed. Besides, we have recognized additional areas, at the moment not foreseen by the EOSC profiles specification, that could further support the process of identification of resources with unique value.

Based on these criteria, up to now, we have selected 7 repositories, 12 thematic, and 12 generic services that have been registered within the project's catalogue and integrated within the pre-production environment. Out of these, 2 repositories, 4 thematic, and 4 generic services are fully integrated, i.e., fully described in terms of the EOSC profiles specification, i.e., can be approved as on-boarded to the EOSC. The EOSC profiles specification recognizes services (there is a dedicated category) from the pre-production environment, and all these services are integrated with the on-boarded resources as well as between themselves. These services in principle can support external users, so all our pre-production services (7 in total) are on-boarded as well. Therefore, in this deliverable, we report 38 resources registered within the catalogue, out of which 17 can be approved as on-boarded to the EOSC.

2. On-boarding procedure

Resources under the project's consideration are grouped into four major categories: repositories, thematic services, generic services, and services from the pre-production environment. Such an organization introduces hierarchy based on resource functions and relationships, which was discussed in the deliverable D3.1 - Best practices for on-boarding and related policies [2]. In NI4OS-Europe, the same on-boarding procedure is used for resources from all four major categories.

The central tool of the on-boarding procedure is the Agora catalogue [5] [7], the resource portfolio management tool that was deployed as a part of our pre-production environment. We use Agora to store all available information about providers and resources. Information entering is performed mainly by the NI4OS-Europe on-boarding team members in collaboration with resource providers. Therefore, the on-boarding team members are main users of Agora, they have the provider admin role, and they can register new resources and update existing ones. Resource providers are also registered within Agora, but with a different role that allows only modification of information associated with a particular resource. The WP5 task leaders are registered as superadmins, who can define the roles of other users, while all interested people can access the catalogue as observers, which is the default role assigned by the system to a new user. At the moment, we have 36 users registered within the Agora catalogue.

Resources and providers could be on-boarded into the EOSC though individual requests or by integration with the external catalogue. The first approach is useful for individual providers, when the total number of provided resources is relatively small. Typically, this happens when a single provider aims to on-board a few self-hosted resources. In such a case, the main on-boarding steps, information gathering, integration, validation, and publication, are performed with the support of the EOSC on-boarding team. Due to a large number of resources stored within national/regional/project catalogues, such an approach would have a huge overhead since the information gathering, integration, and validation are already performed by national/regional/project support teams. Therefore, EOSC provided a dedicated API, to support the publication of already on-boarded resources by the national/regional/project on-boarding teams.

For some time the NI4OS-Europe Agora team, in collaboration with the EOSC team, works on the integration of the Agora catalogue with the central EOSC catalogue. This integration will establish a trusted channel between these two catalogues, and resources/providers marked as on-boarded within the NI4OS-Europe catalogue will be automatically published into the central EOSC one. Practically, the resource on-boarded to the NI4OS-Europe catalogue is also on-boarded to the EOSC. This implies many technical and procedural requirements that are already satisfied by our catalogue, since it strictly follows the EOSC profiles specifications [6], which define common data models for EOSC entities (at the moment, providers and resources) and related taxonomies (such as technology readiness levels, scientific domains, access types, target users, etc.). This ensures the compatibility with the central EOSC catalogue, i.e., resource and provider information collected by NI4OS-Europe is fully compatible with the EOSC profiles specification. Furthermore, we consider a resource and a provider to be on-boarded only when all information marked as mandatory within the EOSC profiles is collected. We expect that the amount of mandatory

information will increase in the upcoming EOSC profiles versions, so we are also trying to collect all the data anticipated to be included in the data model.

However, there is some NI4OS-Europe-specific information which is, due to its uniqueness, not considered by the EOSC profiles. Examples of this are the NI4OS-Europe top-level categorization of resources (repositories, thematic services, generic services, and services from the pre-production environment), OpenDOAR [8] and OpenAIRE [9] repository IDs, etc. In order to preserve compatibility with EOSC, and to add the project's specificities into the data model, we have extended the EOSC profiles with additional entities that are stored independently from Agora. Therefore, besides the Agora database, we have deployed an additional database that keeps project-specific information only. Practically, information required by the EOSC central catalogue is retrieved directly via Agora APIs, while for the project purposes this information is enriched and merged with the information from the additional database that contains project-specific provider/resource metadata fields. Since both databases expose the data via REST API, this is realized by merging two JSON structures into a single one.

The relation between different databases is illustrated in Figure 1. The Agora database is populated manually by the NI4OS-Europe on-boarding team in collaboration with the resource provider via the NI4OS-Europe Agora user interface [10]. API on top of the Agora database sends information to the central EOSC catalogue via a dedicated API and in this way exposes resources on-boarded by the project through the EOSC catalogue and marketplace.
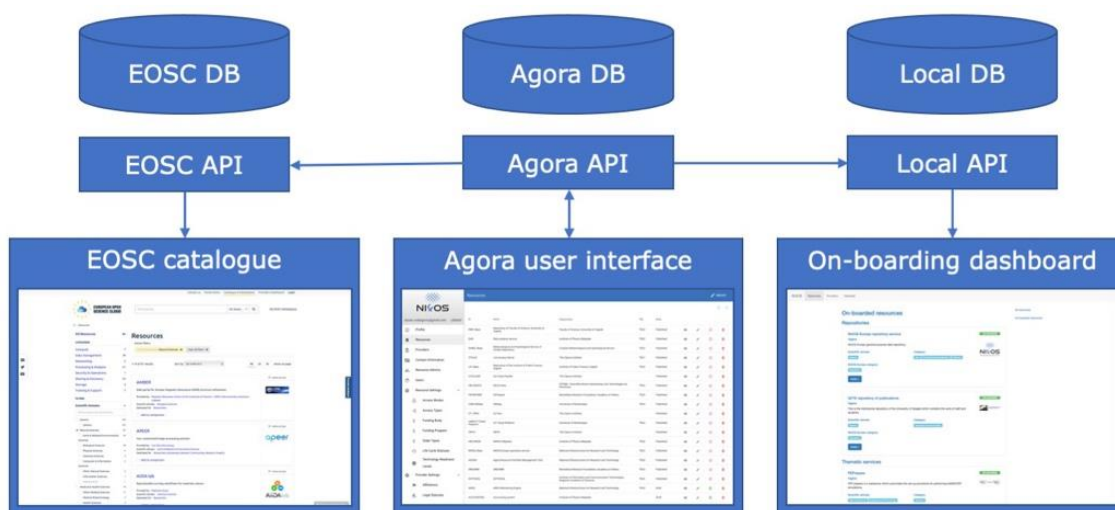
Figure 1: The relation between different databases used during the on-boarding process

On the right-hand side of Figure 1, we have illustrated the project's on-boarding dashboard [11], the platform that combines information stored within the Agora database and the local database, and gives a project-specific view of the registered resources. This tool serves as Agora's front-end but also highlights missing mandatory information per resource, so that the resource provider could track the progress of the on-boarding process. We plan to add many additional features to the tool, but at the moment it lists

on-boarded and candidate resources and providers as illustrated in Figure 2, and gives detailed information about them, as illustrated in Figure 3.

Figure 2: On-boarding dashboard list of resources

Figure 3: On-boarding dashboard detailed resource information

Although in the latest version of the EOSC profiles [6] (version 3.0.0) access mode or criteria for granting access to users is an optional field, we consider this as mandatory fields. From our perspective, it is quite important to clearly define access to a resource in order to be offered through the EOSC federation. Only in this way the potential users can understand under which conditions the resources registered in the catalogue are provided and made available to them. To achieve this, we have also marked as mandatory several other fields from the management section of the EOSC profiles. In particular, the helpdesk page, user manual, terms of use, privacy policy, and access policy are identified as such. This set of management information we consider obligatory, and only resources that provide them can be approved as on-boarded. To support providers, we have prepared guidelines on how to create such policies, as well as templates and useful examples. These are collected at the Wiki page (https://wiki.ni4os.eu/index.php/Policy_templates). The information in the geographical and language availability section are also considered as mandatory. In the project framework, we will on-board only resources open to all European researchers under conditions clearly stated in the policies.

Once the resource is registered within Agora and the support team's e-mail address is supplied, we initiate a helpdesk integration process. The NI4OS-Europe helpdesk [12] is based on the OTRS ticketing system [13], and for each resource we have created an independent support unit. In this way, once the ticket is assigned to a particular support unit, the only member of the resource's support team is notified about a problem via e-mail. Additional, generic support units are created for each resource type (repositories, generic, thematic, and pre-production services), as well as for the first line of support. To submit a question or to report a problem related to resources, the user can open a ticket by sending an e-mail to helpdesk@ni4os.eu e-mail address. The ticket can be also submitted via a dedicated helpdesk form incorporated into the on-boarding dashboard (https://catalogue.ni4os.eu/?_=/helpdesk). Using resource type and resource form fields, the user can specify to which resource the question or problem is related. Otherwise, if not sure, the first line support team will identify the relevant support unit. In both cases, the helpdesk system will automatically inform the customer and support team about ticket creation via e-mail, and record all further communication.

When a support unit is created within the helpdesk, for testing purposes we exchange a test ticket with the support unit members following a pre-defined procedure at the Wiki page (https://wiki.ni4os.eu/index.php/Helpdesk_procedures). After that, we expect all communication with the resource provider to be recorded by the helpdesk system, and we insist on this. At the moment we have 58 tickets (49 closed and 9 open), so 58 different conversations are made via the helpdesk in the last month. The number of tickets created and closed during January 2021 is illustrated in Figure 4. Our aim is not only to collect the total number of tickets but also to try to assess the support's quality in terms of how much time the support unit spent on resolving problems, response times, etc. Therefore, usage of the helpdesk is quite important and can assist in the identification of bottlenecks and suggest ways for provided support improvement.
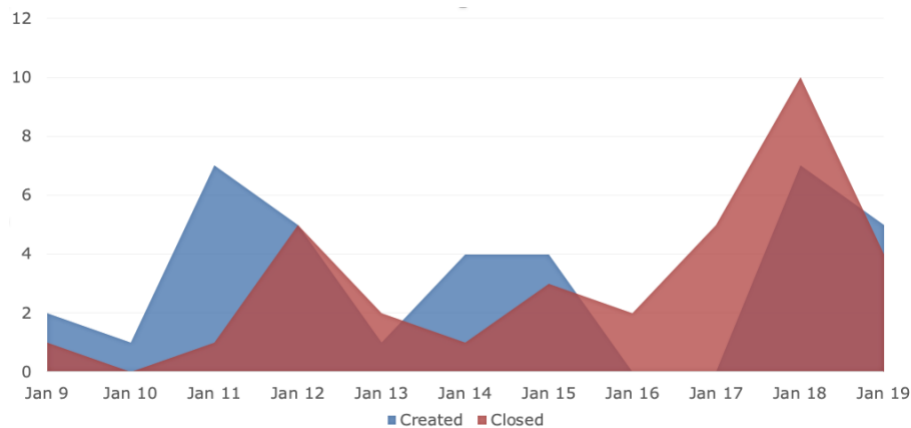
Figure 4: The number of created/closed tickets in January 2021

We also consider as mandatory the integration with the monitoring system during the on-boarding process, and all resources marked as on-boarded are integrated with the Argo monitoring service. At this project phase, we do not insist on the development of resource-specific probes, but rather use an existing basic set of probes to check the availability and reliability of our resources. The on-boarding team does such integration by registering resource end-points into the configuration management database - GOCDB [14]. In the upcoming period, such responsibility will be transferred to resource providers, since the GOCDB is also used for the declaration of downtimes for resource end-points. In parallel, we will work on extracting these data and their presentation on the on-boarding dashboard together with other resource information. In this way, we will systematically ensure that, together with the monitoring information, each resource is equipped with maintenance information, which is also considered in the EOSC profiles schema. GOCDB registry contains information about the topology of our resources, and at the moment, we have registered 21 resources, as illustrated in Figure 5.
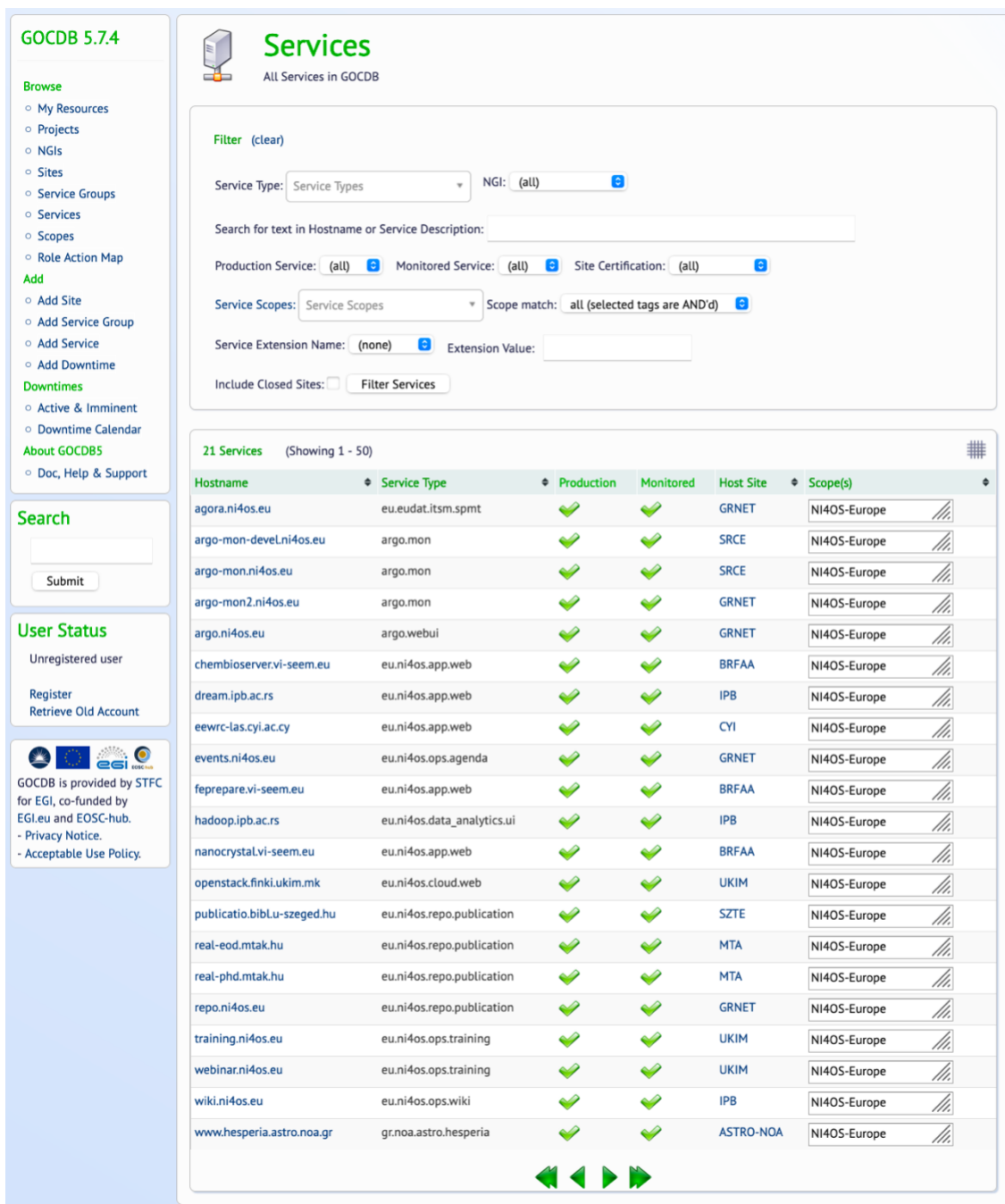
Figure 5: NI4OS-Europe resource registered within the GOCDB

Based on the topology of resources defined within the GOCDB, our Argo monitoring service [15] maps available monitoring probes to resources, and periodically executes the probes. We have the deployed production and development [16] Argo instances, and once a new resource is registered in the GOCDB, it is first tested by the development instance for a verification period of two weeks, and after that, the monitoring is performed by the production instance. According to Argo, the availability of registered resources is very high, around 99%, which is illustrated in Figure 6. We have expected such a high availability since all on-boarded resources have quite a high technology readiness level (TRL 8 or TRL 9). Apart from the basic checks, we would like to create specific metrics

from the user perspective for each resource. Therefore, resource providers are encouraged to develop a resource-specific set of probes, and we have provided detailed guidelines [17] on how to accomplish this. Also, we have gathered more than 100 existing probes and metrics within the Argo poem library [18], whose search mechanism allows resource providers to examine existing probes and potentially identify ones that could be reused for monitoring of their resource.
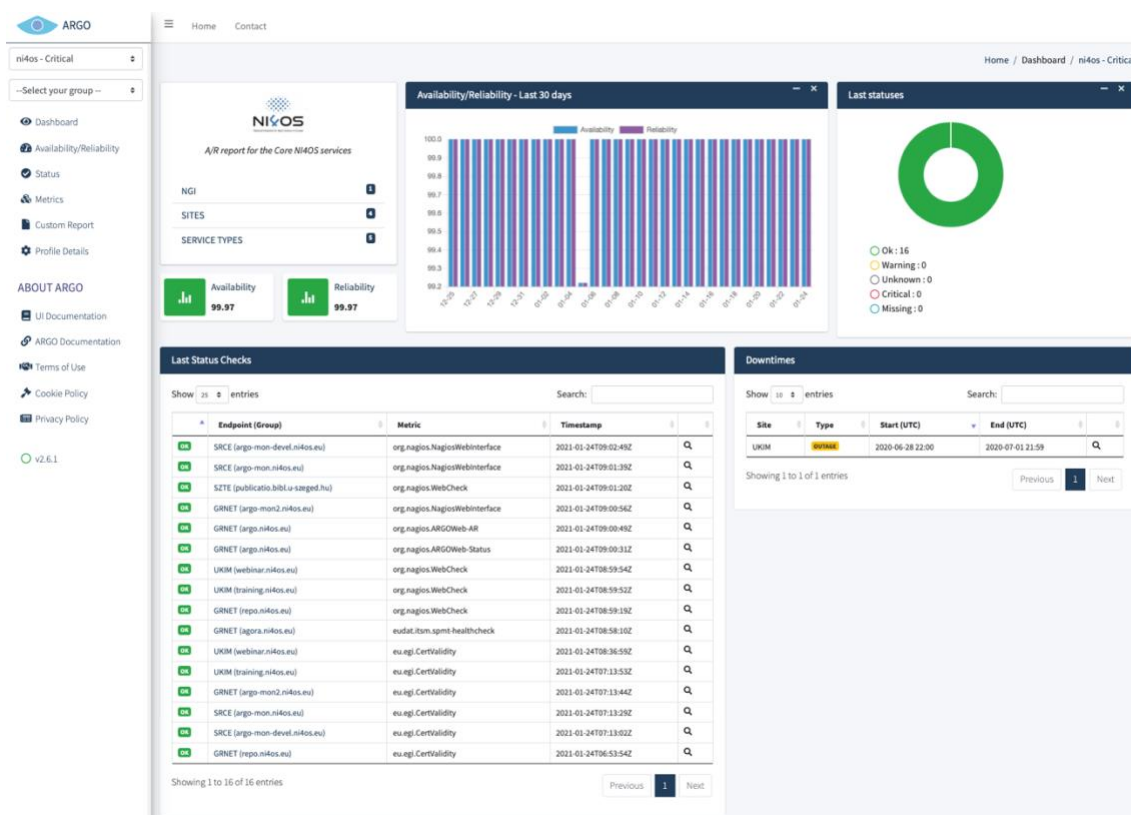
Figure 6: Availability and reliability of NI4OS-Europe on-boarded resources

Although the accounting portal [19] is provided as a part of the NI4OS-Europe pre-production environment, it is currently integrated only with generic HPC/Grid resources, since a plugin that automatically parses HPC/HTC batch system (Torque/Moab or SLURM) log files is available. For other types of resources, it is necessary to develop specific plugins to publish accounting information via REST API or to upload accounting data to the portal manually. In the coming period, once the resources are offered to end-users, and when some usage information is produced, we will be more focused on the integration between resources and the accounting system.

3. On-boarded providers

The provider is defined [6] as an EOSC system user responsible for the provisioning of one or more resources to the EOSC. All partners within the NI4OS-Europe consortium are resource providers, so they are registered within the Agora. In the case of providers, the on-boarding team aims to fully describe, in terms of EOSC profiles mandatory information, providers which resources are also on-boarded. Practically, this means to collect the following EOSC profiles information:

- Basic information
    - ID, a persistent identifier, a unique reference to the provider.
    - Full name and abbreviation of the provider offering the resource and acting as main contact point.
    - Web page with additional information about the provider.
    - Legal status of the provider, which is noted in the registration act/statutes, or in the case of embedded providers, the legal status of the hosting legal entity.
- Classification information
    - Scientific domain, a named group of providers that offer access to the same type of resource or capabilities.
    - Scientific subdomain, a named group of providers that offer access to the same type of resource or capabilities, within the defined domain.
    - Tags associated to the provider to simplify search by relevant keywords.
- Location information
    - Street and number of incorporation or physical location of the provider.
    - Postal code of incorporation or physical location of the provider.
    - City of incorporation or physical location of the provider
    - Country of incorporation or physical location of the provider.
- Marketing information
    - Short description of the provider.
    - Link to the logo/visual identity of the provider.
    - Link to video, slideshow, photos, screenshots with details of the provider.
- Maturity information, the current status of the provider life cycle, as well as list of certifications obtained for the provider (including the certification body and any certificate number), if any.
- Contact information and details about the main contact person, provider manager, and public contact, which include name, email, phone, and position.

Some of these attributes are pre-defined in the form of a list of controlled values, while the others are strings whose maximum length is specified. Also, some attributes may have multiple values. For example, the provider's scientific domain may have several values from a list of controlled values. Verification of entered values according to the EOSC profiles specification is performed by the Agora user interface. Inside the user interface, contact information is separately (in a different section of the user interface) entered and then linked to the provider within the provider's section. Although Agora supports the extension of pre-defined controlled values, such as scientific domains, access types, access modes, etc. we didn't notice any necessity for a new value definition. Values covered by the EOSC profiles are quite comprehensive. An overview of collected provider's information at the NI4OS-Europe on-boarding dashboard is illustrated in Figure 7.

Figure 7: Provider description at the on-boarding dashboard portal

At the moment we have registered 25 resource providers within the catalogue, out of which 6 of them could be approved as on-boarded, i.e., we have collected all the EOSC profiles mandatory information about them. These resource providers are listed in the Table 1, together with the links to the on-boarding dashboard page that gives additional details about the provider.

Provider	On-boarding dashboard page
ATHINA - Erevnitiko Kentro Kainotomias stis Technologies tis Pliroforias	https://catalogue.ni4os.eu/?_=/provider/id/8fdfa222-b3eb-49f4-8387-345e5761d1a2
Biomedical Research Foundation, Academy of Athens	https://catalogue.ni4os.eu/?_=/provider/id/34832a8f-bf05-487a-b00e-fed97e9462ef
National Infrastructures for Research and Technology	https://catalogue.ni4os.eu/?_=/provider/id/94e73408-6b7a-413e-8433-0a8bac4698d0
Institute of Information and Communication Technologies, Bulgarian Academy of Sciences	https://catalogue.ni4os.eu/?_=/provider/id/1e6b2cda-c3fc-45a5-ba51-19f3fbdbaa38
Institute of Physics Belgrade	https://catalogue.ni4os.eu/?_=/provider/id/69a6992d-9a78-44e9-b860-b8947746804c
University of Szeged	https://catalogue.ni4os.eu/?_=/provider/id/a6934df8-3265-4278-93cc-b389c977617b
University Ss. Cyril and Methodius, Faculty of Computer Science and Engineering	https://catalogue.ni4os.eu/?_=/provider/id/afd8e906-d869-4900-aef2-d645364a812d

Table 1: NI4OS-Europe on-boarded providers

4. On-boarded resources

The resource is defined [6] as an asset made available by means of the EOSC system and according to the EOSC rules of participation to EOSC end-users to perform a process useful to deliver value in the context of the EOSC. In the project's context, resources include repositories, thematic services, generic services, and services from the pre-production environment. All of them are on-boarded using the same procedure and the singe Agora catalogue. In terms of EOSC profiles mandatory information each resource to be approved as on-boarded has to provide the following information:

- Basic information
    - ID, a persistent identifier, a unique reference to the resource.
    - Brief and descriptive name of a resource.
    - The name of the organization that manages or delivers the resource, or that coordinates resource delivery in the EOSC environment.
    - The name of the provider that manages or delivers the resource in the EOSC environment.
    - Web page with additional information about the resource.
- Marketing information
    - Short, high-level description in fairly non-technical terms of a) what the resource does and the functionality it provides; b) the benefit to a user/customer delivered by a resource; c) list of customers, communities, users, etc. using the resource.
    - Short catch-phrase for marketing and advertising purposes.
    - Link to the logo/visual identity of the resource.
    - Link to video, slideshow, photos, screenshots with details of the resource.
- Classification information
    - Scientific domain, the branch of science, scientific discipline that is related to the resource.
    - Scientific subdomain, the subbranch of science, scientific subdiscipline that is related to the resource.
    - Category, a named group of resources that offer access to the same type of capabilities.
    - Subcategory, a named group of resources that offer access to the same type of capabilities, within the defined category.
    - Type of users/customers that commissions a provider to deliver a resource.
    - The way a user can access the resource (remote, physical, virtual, etc.).
    - Eligibility/criteria for granting access to users (excellence-based, free-conditionally, free, etc.).
    - Tags associated to the provider to simplify search by relevant keywords.
- Geographical and language availability information. As we said already, the NI4OS-Europe project will on-board only resources available to all European researchers under conditions clearly defined by the resource access policy.
- Resource geographic location, list of geographic locations where data, samples, etc., are stored and processed.
- Contact information and details about the main contact person, provider manager, and public contact, which include name, email, phone, and position.
- Other Contacts

- o Helpdesk e-mail, the email to ask more information from the provider about this resource. This e-mail address is used during the creation of the support unit in the project's helpdesk.
  - o Security contact e-mail, the email to contact the provider for critical security issues about the resource.
- Maturity information in terms of technology readiness level.
- Management information
  - o Helpdesk page, the URL to the helpdesk form of the dashboard portal.
  - o User manual, link to the resource user manual and documentation.
  - o Terms of use, web page describing the rules, resource conditions, and usage policy which one must agree to abide by in order to use the resource.
  - o Privacy policy, link to the privacy policy applicable to the resource.
  - o Access policy, information about the access policies that apply.
  - o Training information, web page on the project's training portal with resource training information.
  - o Status monitoring, resource monitoring information on the project's Argo portal.

Management information ensures practical implementation of various rules of participation, and in the previous period, we tried to emphasize their importance. Therefore, we have organized train-the-trainers event, as well as a self-paced course on ITSM/FISM, so we expect the providers to implement those standards/best practices and increase their maturity with regards to service management.

- Location information
  - o Street and number of incorporation or physical location of the provider.
  - o Postal code of incorporation or physical location of the provider.
  - o City of incorporation or physical location of the provider
  - o Country of incorporation or physical location of the provider.
- Contact information and details about the main contact person, provider manager, and public contact, which include name, email, phone number, and position.

The Agora user interface [10] has different forms for provider and resource data entering. This feature allows us to easily link the same provider with several resources, which is very useful when the same provider offers several resources. Resource organization and resource providers in the resource data model are linked to the organizations described by the provider data model. Practically, the user interface lists all registered providers, so the resource manager can perform linking.

At the moment the resource maturity level is assessed only by the technology readiness level, and the level of details will be increased in this section by a specification of used open-source technologies, last update of the service, and summary of the resource features updated from the previous version. This applies also to dependencies information, which will link resources between themselves within the catalogue. For example, DREAMM thematic service (described in Section 4.2.2) runs on top of Avitohol cloud generic service (listed in Table 10), and we would like to store this type of information within the catalogue as well. In addition, the current level of resource description will be enriched with the attribution information, such as the name of the funding body/program/project that supported the development and operation. Although it is clear that resources published

within Agora and displayed at the dashboard portal are on-boarded within the framework of the NI4OS-Europe project, once are moved to the central EOSC catalogue this might not be so visible. Therefore, based on the attribution data such information could easily propagate to the central catalogue and enable grouping of resources per project if necessary.

4.1. On-boarded repositories

In the previous period, we have successfully on-boarded, i.e., described in terms of EOSC resource profiles specification two repositories: NI4OS-Europe repository service provided and by the National Infrastructures for Research and Technology, Greece, and SZTE repository of publications provided and hosted by the University of Szeged, Hungary. In addition to these, six other repositories are registered within the catalogue, but some aspects of resource description are still under development. Details about on-boarded repositories are given in the subsections, while the candidates are listed in Table 2 together with the links to the resource on-boarding dashboard page where the up-to-dated details are provided.

Repository	On-boarding dashboard page
Repository of Faculty of Science, University of Zagreb	https://catalogue.ni4os.eu/?_=/resource/id/0eed0440-917e-44a9-a324-aee1d0873744
Meteorological and Hydrological Service of Croatia Repository	https://catalogue.ni4os.eu/?_=/resource/id/33c74095-a1e1-4326-9511-77a794a39489
Repository of the Institute of Public Finance, Zagreb	https://catalogue.ni4os.eu/?_=/resource/id/3d3a3b00-0936-44f6-b01a-bfcda9b0145d
HELIX Data	https://catalogue.ni4os.eu/?_=/resource/id/4c7896d0-0945-42d0-bdf7-c31a090f00ae
UKIM Repository	https://catalogue.ni4os.eu/?_=/resource/id/ee7096b3-6aa7-468e-accd-e6dcad722678

Table 2: On-boarding repository candidates registered within the Agora catalogue

4.1.1.   NI4OS-Europe repository service

NI4OS Repository Service (NRS) is the main storage service of a community that holds "Regional Community Datasets". The NRS is also the platform to host all kinds of additional data such as publications (and their associated data), software (or references to software), workflow descriptions (e.g. how to generate research data) or even materials targeting the general public (e.g. images, videos etc.). NRS is integrated with a persistent identifier service as an assigned PID is required for each digital object (item, collection, community).

The repository service allows NI4OS-Europe users to deposit and share data via a user-friendly web interface. It can host publications and their associated data or software. It

automatically generates a Persistent Identifier for each shared item. Access to shared items can be public or limited to selected repository users.

Information we have collected about the NI4OS-Europe repository service is summarized in Table 3.

Resource organization	National Infrastructures for Research and Technology
Resource provider	National Infrastructures for Research and Technology
Web page	https://repo.ni4os.eu/
OpenDOAR	http://v2.sherpa.ac.uk/id/repository/9937
OpenAIRE	https://explore.openaire.eu/search/dataprovider?datasourceId=opendoar::ad80947c9909dd9d70739ca2b8f3fd2d
Scientific domain	Generic
Scientific subdomain	Generic
Category	Data, Scholarly Communication, Software
Subcategory	Data, Discovery, Publication
Access type	Virtual
Access mode	Free
Helpdesk page	https://catalogue.ni4os.eu/?_=/helpdesk/Repositories/NI4OS-Europe%20repository%20service
Privacy policy	https://repo.ni4os.eu/privacy
Status monitoring	https://argo.ni4os.eu/ni4os/report-status/Critical/SITES/GRNET/eu.ni4os.repo.publication
Geographic location	Greece
Technology readiness level	TRL9
More information	https://catalogue.ni4os.eu/?_=/resource/id/b35269cb-82f2-4aec-bc61-f2684069bff3

Table 3: NI4OS-Europe repository service information

4.1.2.   SZTE repository of publications

The intention of the SZTE Repository of Publications (Table 4) is to make the full text of publications created as a result of scientific and artistic activities at the University available for the widest possible academic audience. Depositing works at the repository secures their long term archiving, and can also increase their visibility and number of citations. This latter is also because uploaded documents are indexed by general search engines (e.g. Google, Google Scholar) and professional databases (e.g. BASE, MTA OAI).

Resource organization	University of Szeged
Resource provider	University of Szeged
Web page	http://publicatio.bibl.u-szeged.hu/

OpenDOAR	http://v2.sherpa.ac.uk/id/repository/2868
OpenAIRE	https://explore.openaire.eu/search/dataprovider?datasourceId=opendoar_::8b3bac12926cc1d9fb5d68783376971d
Scientific domain	Generic
Scientific subdomain	Generic
Category	Scholarly Communication
Subcategory	Discovery, Publication
Access type	Virtual
Access mode	Free
Helpdesk page	https://catalogue.ni4os.eu/?_=/helpdesk/Repositories/SZTE%20repository%20of%20publications
Terms of use	http://publicatio.bibl.u-szeged.hu/policies.html
Privacy policy	http://publicatio.bibl.u-szeged.hu/policies.html
Access policy	http://publicatio.bibl.u-szeged.hu/policies.html
Status monitoring	https://argo.ni4os.eu/ni4os/report-status/Critical/SITES/SZTE/eu.ni4os.repo.publication
Geographic location	Hungary
Technology readiness level	TRL9
More information	https://catalogue.ni4os.eu/?_=/resource/id/db7e9222-8c75-4548-a862-b464f5991210

Table 4: SZTE repository of publications information

4.2. On-boarded thematic services

Within the Agora catalogue, we have registered 12 thematic services, 4 of them are marked as on-boarded, while the other 8 as candidates. In this section, we describe on-boarded thematic services, while the candidates together with the links to the on-boarding dashboard page are listed in Table 5. In the previous period, our main goal was to equip the services with the appropriate user manual, to perform integration with the project's helpdesk, and to produce relevant policies.

Repository	On-boarding dashboard page
Live Access Server	https://catalogue.ni4os.eu/?_=/resource/id/3d243e51-9224-4ae3-810b-9d3c5a311193
Clowder4DCH	https://catalogue.ni4os.eu/?_=/resource/id/d0c2baae-45dc-496f-922c-dcb051ce5fc0
DREAM	https://catalogue.ni4os.eu/?_=/resource/id/4799d746-e9cd-4a61-be59-2e0a2aa88600
Schrodinger	https://catalogue.ni4os.eu/?_=/resource/id/c28834c3-0402-421a-8cd3-be8f862f7b10

Airpolution prediction	https://catalogue.ni4os.eu/?_=/resource/id/ce2f1db4-c0d2-436c-bb45-ba9ca04fb625
OMApp	https://catalogue.ni4os.eu/?_=/resource/id/7919d83d-a02d-4b39-87cb-a1db40419f06
IoT Cloud Platform	https://catalogue.ni4os.eu/?_=/resource/id/923fb0a9-253f-4ddc-a71c-3fb1dbf017c7
High Energy Solar Particle Events foRecastIng and Analysis	https://catalogue.ni4os.eu/?_=/resource/id/c41e1ab7-bf1e-488b-bb00-c0edc0b844e2

Table 5: On-boarding thematic service candidates registered within the Agora catalogue

4.2.1.   FEPrepare

FEP prepare is a web server, which automates the set-up procedure for performing NAMD/FEP simulations. Automating free energy perturbation calculations is a step forward to delivering high throughput calculations for accurate predictions of relative binding affinities before a compound is synthesized, and consequently save enormous time and cost.

Information we have collected about the FEPrepare thematic service is summarized in Table 6.

Resource organization	Biomedical Research Foundation, Academy of Athens
Resource provider	National Infrastructures for Research and Technology
Web page	https://feprepare.vi-seem.eu/
Scientific domain	Natural Sciences, Engineering & Technology
Scientific subdomain	Chemical sciences, Chemical engineering
Category	Software
Subcategory	Platform
Access type	Virtual
Access mode	Free
Helpdesk page	https://catalogue.ni4os.eu/?_=/helpdesk/Thematic%20services/FEPrepare
User manual	https://catalogue.ni4os.eu/pdf/FEPrepare-User_manual.pdf
Terms of use	https://catalogue.ni4os.eu/pdf/FEPrepare-Terms_of_use.pdf
Privacy policy	https://catalogue.ni4os.eu/pdf/FEPrepare-Privacy_policy.pdf
Access policy	https://catalogue.ni4os.eu/pdf/FEPrepare-Acceptable_use_policy.pdf
Status monitoring	https://argo-devel.ni4os.eu/ni4os/report-status/Critical/SITES/BRFAA/eu.ni4os.app.web/feprepare.vi-seem.eu
Geographic location	Greece
Technology readiness level	TRL9

| More information | https://catalogue.ni4os.eu/?_=/resource/id/4d2f5c58-d32d-46ce-b075-8b278889fe96 |

Table 6: FEPrepare thematic service information

4.2.2. DREAMM

DREAMM service, which description is summarized in Table 7, is a novel machine learning tool that predicts the protein-membrane interfaces of peripheral membrane protein, and optionally predicts binding sites near the predicted membrane-penetrating residues in protein conformational ensembles. As an output, the user can retrieve the membrane-penetrating residues in a .csv file and if the user's choice was to predict binding sites, a .zip file will be downloaded including the abovementioned .csv file, the binding pocket predictions, the visualizations, and summarized the binding site clustering results.

Resource organization	Biomedical Research Foundation, Academy of Athens
Resource provider	Biomedical Research Foundation, Academy of Athens
Web page	https://dreamm.ni4os.eu/
Scientific domain	Natural Sciences
Scientific subdomain	Physical sciences, Biological sciences, Computer and information sciences
Category	Software
Subcategory	Platform
Access type	Virtual
Access mode	Free
Helpdesk page	https://catalogue.ni4os.eu/?_=/helpdesk/Thematic%20services/DREAMM
User manual	https://catalogue.ni4os.eu/pdf/DREAMM-User_manual.pdf
Terms of use	https://catalogue.ni4os.eu/pdf/DREAMM-Terms_of_use.pdf
Privacy policy	https://catalogue.ni4os.eu/pdf/DREAMM-Privacy_policy.pdf
Access policy	https://catalogue.ni4os.eu/pdf/DREAM-Acceptable_use_policy.pdf
Geographic location	Greece
Technology readiness level	TRL9
More information	https://catalogue.ni4os.eu/?_=/resource/id/b35269cb-82f2-4aec-bc61-f2684069bff3

Table 7: DREAMM thematic service information

4.2.3. NanoCrystal

NanoCrystal is a novel web-based crystallographic tool that creates nanoparticle models from any crystal structure guided by their preferred equilibrium shape under standard conditions according to the Wulff morphology (crystal habit). Users can upload a cif file, define the Miller indices and their corresponding minimum surface energies according to

the Wulff construction of a particular crystal, and specify the size of the nanocrystal. As a result, the nanoparticle is constructed and visualized, and the coordinates of the atoms are output to the user.

NanoCrystal thematic service description is summarized in Table 8.

Resource organization	Biomedical Research Foundation, Academy of Athens
Resource provider	National Infrastructures for Research and Technology
Web page	https://nanocrystal.vi-seem.eu/
Scientific domain	Engineering & Technology
Scientific subdomain	Materials engineering, Nano-technology
Category	Software
Subcategory	Platform
Access type	Virtual
Access mode	Free
Helpdesk page	https://catalogue.ni4os.eu/?_=/helpdesk/Thematic%20services/NanoCrystal
User manual	https://catalogue.ni4os.eu/pdf/NanoCrystal-User_Manual.pdf
Terms of use	https://catalogue.ni4os.eu/pdf/NanoCrystal-Terms_of_use.pdf
Privacy policy	https://catalogue.ni4os.eu/pdf/NanoCrystal-Privacy_policy.pdf
Access policy	https://catalogue.ni4os.eu/pdf/NanoCrystal-Acceptable_use_policy.pdf
Training information	https://training.ni4os.eu/course/view.php?id=56
Status monitoring	https://argo-devel.ni4os.eu/ni4os/report-status/Critical/SITES/BRFAA/eu.ni4os.app.web/nanocrystal.vi-seem.eu
Geographic location	Greece
Technology readiness level	TRL9
More information	https://catalogue.ni4os.eu/?_=/resource/id/c1f8d7fb-4250-4278-ba8d-4d1df4a85df5

Table 8: NanoCrystal thematic service information

4.2.4.   ChemBioServer

ChemBioServer (description summary is given in Table 9) is a web server for filtering, clustering and networking of chemical compound libraries facilitating both drug discovery and repurposing. It provides researchers the ability to (i) browse and visualize compounds along with their physicochemical and toxicity properties; (ii) perform property-based filtering of chemical compounds; (iii) explore compound libraries for lead optimization based on perfect match substructure search; (iv) re-rank virtual screening results to achieve selectivity for a protein of interest against different protein members of the same family, selecting only those compounds that score high for the protein of interest; (v) perform clustering among the compounds based on their physicochemical properties providing representative compounds for each cluster; (vi) construct and visualize a structural similarity network of compounds providing a set of network analysis metrics; (vii) combine a given set of compounds with a reference set of compounds into a single

structural similarity network providing the opportunity to infer drug repurposing due to transitivity; (viii) remove compounds from a network based on their similarity with unwanted substances (e.g. failed drugs) and (ix) build custom compound mining pipelines.

Resource organization	Biomedical Research Foundation, Academy of Athens
Resource provider	National Infrastructures for Research and Technology
Web page	https://chembioserver.vi-seem.eu/
Scientific domain	Natural Sciences
Scientific subdomain	Chemical sciences
Category	Software
Subcategory	Platform
Access type	Virtual
Access mode	Free
Helpdesk page	https://catalogue.ni4os.eu/?_=/helpdesk/Thematic%20services/ChemBioServer
User manual	https://catalogue.ni4os.eu/pdf/ChemBioServer-User_manual.pdf
Terms of use	https://catalogue.ni4os.eu/pdf/ChemBioServer-Terms_of_use.pdf
Privacy policy	https://catalogue.ni4os.eu/pdf/ChemBioServer-Privacy_policy.pdf
Access policy	https://catalogue.ni4os.eu/pdf/ChemBioServer-Acceptable_use_policy.pdf
Status monitoring	https://argo-devel.ni4os.eu/ni4os/report-status/Critical/SITES/BRFAA/eu.ni4os.app.web/chembioserver.vi-seem.eu
Geographic location	Greece
Technology readiness level	TRL9
More information	https://catalogue.ni4os.eu/?_=/resource/id/d06512f9-b287-4452-8a2f-5742eccb44c4

Table 9: ChemBioServer thematic service information

4.3. On-boarded generic services

We have fully described in terms of EOSC profiles specification 4 generic services, while an additional 8 are partially described. So, in total 12 generic services are registered within the project's catalogue. Table 10 lists partially on-boarded services, while details on the fully on-boarded services are presented in this section.

Repository	On-boarding dashboard page
Simple storage service	https://catalogue.ni4os.eu/?_=/resource/id/986f7994-e45a-4f07-b05f-df40f78a8508
Data discovery service	https://catalogue.ni4os.eu/?_=/resource/id/117c2e24-c2af-47b4-8dcc-7a2892d88e84

Avitohol cloud	https://catalogue.ni4os.eu/?_=/resource/id/45ba1dd6-6bc5-4974-b20c-c27dbbb6aa6f
ARIS	https://catalogue.ni4os.eu/?_=/resource/id/fe5d5668-e435-4bae-ac93-b44c7e04f155
Archival service	https://catalogue.ni4os.eu/?_=/resource/id/a40c074d-f507-48b8-9970-8276bc5b0c87
CyI Cloud Facility	https://catalogue.ni4os.eu/?_=/resource/id/43b7bd0a-39e1-47d8-807a-5cf5554ece2e
Cy-Tera	https://catalogue.ni4os.eu/?_=/resource/id/7ef76b56-0fbc-4bb6-883a-127539b535d6
ONYX	https://catalogue.ni4os.eu/?_=/resource/id/9695ac3f-8169-4646-be7e-372778dec0e5

Table 10: On-boarding generic service candidates registered within the Agora catalogue

4.3.1. AVITOHOL

The supercomputer Avitohol was at 331st place in the TOP 500 list of supercomputers. It is built with HP Cluster Platform SL250S GEN8 (150 servers), Intel Xeon E5-2650 v2 8C 2.6GHz CPUs (300 CPUs), non-blocking InfiniBand FDR, 300 Intel Xeon Phi 7120P co-processors. It provides 412 TFlops of performance for diverse scientific and industrial applications. Users from science and industry with substantial computational needs use it to achieve their results faster and to solve bigger problems that are beyond the reach of ordinary clusters.

A summarized description of the Avitohol in terms of EOSC profiles specification is given in the Table 11.

Resource organization	Institute of Information and Communication Technologies, Bulgarian Academy of Sciences
Resource provider	Institute of Information and Communication Technologies, Bulgarian Academy of Sciences
Web page	http://www.hpc.acad.bg/avitohol/
Scientific domain	Generic
Scientific subdomain	Generic
Category	Compute
Subcategory	Job Execution
Access type	Remote
Access mode	Peer-reviewed
Helpdesk page	http://www.hpc.acad.bg/helpdesk/
User manual	http://hpc.acad.bg/avitohol/AVITOHOL_Best_Practice_Guide.pdf
Terms of use	http://www.hpc.acad.bg/access-policies/AVITOHOL_AUP.pdf
Privacy policy	http://www.hpc.acad.bg/access-policies/AVITOHOL_privacy.pdf
Access policy	http://www.hpc.acad.bg/access-policies/AVITOHOL_AUP.pdf

Geographic location	Bulgaria
Technology readiness level	TRL9
More information	https://catalogue.ni4os.eu/?_=/resource/id/bd3651fb-3a04-4521-b1b7-79abcd1292e2

Table 11: Avitohol generic service information

4.3.2.　　PARADOX-IV cluster

PARADOX-IV cluster represents the fourth major upgrade of the PARADOX cluster and became operational during September 2013. The cluster consists of 106 working nodes and 3 service nodes. Working nodes (HP ProLiant SL250s Gen8, 2U height) are configured with two Intel Xeon E5-2670 8-core Sandy Bridge processors, at a frequency of 2.6 GHz and 32 GB of RAM (2 GB per CPU-core). The total number of new processor-cores in the cluster is 1696. Each working node contains an additional GP-GPU card (NVIDIA Tesla M2090) with 6 GB of RAM. With a total of 106 NVIDIA Tesla M2090 graphics cards, PARADOX is a premier computer resource in the wider region, which provides access to a large production GPU cluster and new technology. The peak computing power of PARADOX is 105 TFlops. Other technical information is provided in the table below.

One service node (HP DL380p Gen8), equipped with an uplink of 10 Gbps, is dedicated to cluster management and user access (gateway machine). All cluster nodes are interconnected via Infiniband QDR technology, through a non-blocking 144-port Mellanox QDR Infiniband switch. The communication speed of all nodes is 40 Gbps in both directions, which is a qualitative step forward over the previous (Gigabit Ethernet) PARADOX installation. The administration of the cluster is enabled by an independent network connection through the iLO (Integrated Lights-Out) interface integrated on motherboards of all nodes.

PARADOX cluster is installed in four water-cooled racks. The cooling system consists of 4 cooling modules (one within each rack), which are connected via a system of pipes with a large industrial chiller and configured so as to minimize power consumption.

PARADOX-IV cluster description is summarized in Table 12.

Resource organization	Institute of Physics Belgrade
Resource provider	Institute of Physics Belgrade
Web page	https://www.scl.rs/PARADOX_User_Guide/p4-user-guide.html
Scientific domain	Natural Sciences, Engineering & Technolog
Scientific subdomain	Chemical sciences, Physical sciences, Biological sciences, Environmental biotechnology, Mathematics, Other natural sciences, Materials engineering, Chemical engineering, Computer and information sciences, Environmental engineering, Industrial biotechnology, Nano-technology, Mechanical engineering, Earth and related environmental sciences
Category	Compute
Subcategory	Job Execution
Access type	Virtual

Access mode	Peer-reviewed
Helpdesk page	https://catalogue.ni4os.eu/?_=/helpdesk/Generic%20services/PARADOX-IV%20cluster
User manual	https://www.scl.rs/PARADOX_User_Guide/p4-user-guide.html
Terms of use	https://www.scl.rs/PARADOX_User_Guide/policies/terms-of-use.html
Privacy policy	https://www.scl.rs/PARADOX_User_Guide/policies/privacy-policy.html
Access policy	https://www.scl.rs/PARADOX_User_Guide/policies/acceptable-use-policy.html
Geographic location	Serbia
Technology readiness level	TRL9
More information	https://catalogue.ni4os.eu/?_=/resource/id/f3d0551b-1802-4635-bba1-7acfc67ed37d

Table 12: PARADOX-IV cluster information

4.3.3.   Data analysis service

Data analysis service (description summarized in Table 13) or PARADOX Hadoop cluster consists of a single name node that runs the YARN resource manager, and three additional data nodes. The name node is hosted on a machine with 4-core Intel Xeon E3-1220v3 CPU running at 3.1 GHz, with 4 GB of RAM, and 500 GB of local hard disk storage. Each of the data nodes, which perform the computation and storage, are hosted on machines with 24-core Intel Xeon E5-2620 CPUs at 2.4 GHz, with 64 GB of RAM and 2 TB of storage. In total, the cluster provides access to 60 CPU cores, 180 GB of RAM and 5.3 TB of storage in HDFS.

In the analysis of very large datasets, the movement of data can be present a far more severe bottleneck than the actual computation. Therefore, the PARADOX Hadoop cluster is designed to overlap computation and data storage operations, i.e., to enable performing of computation on the same machine(s) that store the corresponding data.

Resource organization	Institute of Physics Belgrade
Resource provider	Institute of Physics Belgrade
Web page	https://www.scl.rs/PARADOX_User_Guide/hadoop-user-guide.html
Scientific domain	Natural Sciences, Engineering & Technolog
Scientific subdomain	Chemical sciences, Physical sciences, Biological sciences, Environmental biotechnology, Mathematics, Other natural sciences, Materials engineering, Chemical engineering, Computer and information sciences, Environmental engineering, Industrial biotechnology, Nano-technology, Mechanical engineering, Earth and related environmental sciences
Category	Data analysis, Compute
Subcategory	2D/3D Digitisation, Job Execution, Machine Learning, Image/data analysis, Forecast, Data extrapolation, Visualization, Artificial Intelligence, Workfows
Access type	Virtual
Access mode	Peer-reviewed

Helpdesk page	https://ni4os.ipb.ac.rs/?_=/helpdesk/Generic%20services/Data%20analysis%20service
User manual	https://www.scl.rs/PARADOX_User_Guide/hadoop-user-guide.html
Terms of use	https://www.scl.rs/PARADOX_User_Guide/policies/terms-of-use.html
Privacy policy	https://www.scl.rs/PARADOX_User_Guide/policies/privacy-policy.html
Access policy	https://www.scl.rs/PARADOX_User_Guide/policies/acceptable-use-policy.html
Geographic location	Serbia
Technology readiness level	TRL9
More information	https://catalogue.ni4os.eu/?_=/resource/id/0fc83dce-9a1f-4aed-b431-85ec5cd961ef

Table 13: Data analysis service information

4.3.4. FINKI cloud

Openstack cloud is deployed at the Faculty of Computer Science and Engineering, UKIM. Cloud infrastructure is based on Openstack and is hosted on 15 Huawei servers, each with 128GB RAM and 20 HT CPU cores, which makes in total 300 vCPU cores and 37TB SSD and 32 TB SAS storage. The system is in production since 2017 as a National cloud system. The connectivity to the Internet is 1Gbit through MARNET provided link to GEANT. Currently, the system hosts templates for all popular Linux distributions and Windows variations. The primary target communities are the ones from the long tail of science.

The infrastructure is fully integrated with the eduGAIN/NI4OS AAI, enabling seamless access to virtualize computing, network, and storage resources.

A summarized description of the FINKI cloud resource in terms of EOSC profiles specification is given in the Table 14.

Resource organization	University Ss. Cyril and Methodius, Faculty of Computer Science and Engineering
Resource provider	University Ss. Cyril and Methodius, Faculty of Computer Science and Engineering
Web page	https://openstack.finki.ukim.mk/
Scientific domain	Generic
Scientific subdomain	Generic
Category	Compute
Subcategory	Virtual Machine Management
Access type	Virtual
Access mode	Free
Helpdesk page	https://catalogue.ni4os.eu/?_=/helpdesk/Generic%20services/FINKI%20cloud
User manual	https://openstack.finki.ukim.mk/assets/user-guide.html
Terms of use	https://openstack.finki.ukim.mk/assets/terms-of-use.html

Privacy policy	https://openstack.finki.ukim.mk/assets/privacy-policy.html
Access policy	https://openstack.finki.ukim.mk/assets/acceptable-use-policy.html
Geographic location	North Macedonia
Technology readiness level	TRL9
More information	https://catalogue.ni4os.eu/?_=/resource/id/e9ba68b9-5f2d-4359-9372-b6fffd8ccb24

Table 14: FINKI cloud information

4.4. On-boarded services from the pre-production environment

As a part of on-boarding activities, we have on-boarded all services from the pre-production environment. All these services are already described in the deliverable D3.2 - First report on pre-production environment, and here in Table 15 we list links to the on-boarding dashboard pages where EOSC profiles description is given.

Repository	On-boarding dashboard page
Agora Resource Portfolio Management Tool	https://catalogue.ni4os.eu/?_=/resource/id/ba714802-5c0c-4401-bba6-033fb5b6b956
Configuration Management Database	https://catalogue.ni4os.eu/?_=/resource/id/fa478456-d9ef-4e65-90a6-252954fcc304
NI4OS Login	https://catalogue.ni4os.eu/?_=/resource/id/e8761c52-d713-4461-8dc9-87445f359f62
Argo monitoring	https://catalogue.ni4os.eu/?_=/resource/id/bfa1a035-d6b5-412d-b1b3-43ab9780a7d1
Accounting system	https://catalogue.ni4os.eu/?_=/resource/id/c1f5e284-15b4-4f53-8c06-9aa5b986226f
Helpdesk	https://catalogue.ni4os.eu/?_=/resource/id/9ba9cd29-42af-44ab-af9f-b799d1518c55
NI4OS-Europe training and webinar portal	https://catalogue.ni4os.eu/?_=/resource/id/c84ac92d-4f1a-4941-8d61-e583a001660e

Table 15: On-boarded services from the pre-production environment

5. Conclusions

During the on-boarding process, it was crucial to identify resources that have an added value for the EOSC community and users. Therefore, first, we performed a categorization of the available resources, in order to distinguish different aspects of the resource development stage, and later to quantify the potential impact of a particular resource with respect to each aspect. This was done by analyzing the information foreseen by the EOSC profiles that are incorporated into the Agora catalogue. As a result, three main technical aspects were distinguished based on the EOSC profiles specification: the development stage, the EOSC integration stage, and the service management levels.

We have also identified additional categories not covered by the EOSC profiles at the moment, but yet relevant for resource assessment. For generic services, it is the capacity of the resources measured in the number of the CPU cores, storage size, amount of GPU cards, or any other unit that will permit comparison of generic services within the catalogue. For thematic services, it is necessary to develop a more complex categorization in terms of scientific relevance, innovation potential, scientific or social impact. This can be only done by a separate review of each thematic service. However, in order to perform such a review, we need to gather additional information, such as recent bibliographic references, size of the community, number of registered patents, international collaborations, etc., which is required. Similarly, for repositories, it might be useful to include a metric to measure the quality of data stored within the repository, for example, the number of empty metadata fields, the total number of entries, etc.

Up to now, we have selected 7 repositories, 12 thematic, and 12 generic services that have been registered within the Agora catalogue and integrated within the pre-production environment. We have fully on-boarded 2 repositories, 4 thematic, and 4 generic services, while the other resources are partially on-boarded. All services from the pre-production environment (7 services in total) were also on-boarded, and therefore the total number of resources available within the catalogue is currently 38.