H2020-INFRAEOSC-2018-3

NI4OS-Europe

National Initiatives for Open Science in Europe

Deliverable D5.4
Second report on provider and repository integration and horizontal service delivery

---

Lead beneficiary(s):	UKIM
Author(s):	NI4OS-Europe consortium
Status –Version:	Final – f
Date:	January 18, 2023
Dissemination Level:	Public

Abstract: Deliverable D5.4 – Second report on provider and repository integration and horizontal service delivery - presents the resource and provider on-boarding status. The deliverable gives details on the updated on-boarding procedure and presents integrated repositories, thematic and generic services, as well as services from the pre-production environment and relevant statistics.

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

COPYRIGHT NOTICE

Document Revision History

Date	Issue	Author/Editor/Contributor	Summary of main changes
December 7, 2022	a	Anastas Mishev, Dušan Vudragović	Initial ToC
January 9, 2023	b	Anastas Mishev, Dušan Vudagović	Initial contentment provided
January 13, 2023	c	Andreas Athenodorou, Svetlozar Yordanov, Themis Zamani, Kostas Koumantaros, Dusan Vudgragovic, Anastas Mishev, Nicolas Liabotis, Ognjen Prnjat	Content provided
January 16	d	Dusan Vudragovic, Anastas Mishev	Content contribution
January 17	e	Ognjen Prnjat, Dusan Vudragovic, Anastas Mishev	Final draft version
January 18	f	Milutin Radonjic	Quality control, final version

Table of contents

References

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

[2]   First report on the provider and repository integration, NI4OS-Europe deliverable D5.2: https://doi.org/10.5281/zenodo.4485708

[3]   EOSC Portal Profiles, https://wiki.eoscfuture.eu/display/PUBLIC/EOSC+Portal+Profiles

[4]   Best practices for on-boarding and related policies 2nd version, NI4OS-Europe deliverable D3.1, https://doi.org/10.5281/zenodo.5078116

[5]   Final report on pre-production environment, NI4OS-Europe deliverable D3.5.

[6]   NI4OS-Europe AGORA user interface, https://agora.ni4os.eu/.

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

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

[9]   NI4OS-Europe ARGO monitoring service, https://argo.ni4os.eu/.

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

[11] EOSC Exchange, https://eoscfuture.eu/ker/eosc-exchange/

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
SEE	South Eastern Europe
SLURM	Simple Linux Utility for Resource Management
SPMT	Service Portfolio Management Tool

Executive summary

What is the focus of this Deliverable?

Building up on the deliverable D5.2 First report on the provider and repository integration [2], this deliverable represents the status of the on-boarded providers and resources approaching the end of the NI4OS-Europe project [1]. It addresses the changes in the metadata for the description of the providers and resources that evolved during the project lifetime as well as the support given to the providers within the pre-production environment. It also contains a list of all on-boarded providers and resources up-to-date.

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

This deliverable is one of the final deliverables toward the end of the project, listing most important achievements in the identification and on-boarding of the resources and the providers from the project partners countries. It also provides basics for the deliverables D6.8 Report on the open call and D7.7 Sustainability report.

What are the deliverable contents?

The deliverable gives an overview of the on-boarding and horizontal integration of resources from the point of view of the pre-production environment, more specifically monitoring, accounting and helpdesk, followed by the latest changes to the EOSC profiles for provides and changes in the EOSC profiles for resources [3]. The updated on-boarding procedure [4] based on the feedback from the providers gathered during the project lifetime is described. The deliverable lists the on-boarded providers, followed by descriptions of all the on-boarded resources, categorized in generic, thematic and repositories, along with the services from the pre-production environment. The usage of the NI4OS-Europe services by the wider research community through the open call is presented next.

Conclusions and recommendations

The NI4OS-Europe project consortium provided the necessary technical environment and policy framework for the on-boarding of regional service providers and resources to EOSC, meeting the proposed KPIs. Continual support for the pre-production environment is crucial for the sustainability and future resource integration from the SEE region.

1. Introduction

The successful on-boarding of the providers and resources from the SEE to the NI4OS-Europe catalog is a result of the well-established technical, policy and human framework. The human support was provided through the on-boarding team, consisting of representatives from all the partner countries. The policy framework is based on the commonly established EOSC profiles for providers and resources, implemented in the service portfolio management tool (SPMT) **Error! Reference source not found.**. Finally, the technical framework is implemented using the pre-production environment [5], enabling the providers set of tools to integrate and validate their resources to be on-boarded to EOSC.

Key elements that provided the service integration of all on-boarded services, beside the SPMT, were the monitoring tool ARGO [9], the in-hose developed accounting system [10], as well as the help-desk system [8] playing a key role for the end-user interaction and support.

The NI4OS-Europe Monitoring service provides a flexible and scalable framework for monitoring status, availability and reliability of a wide range of services provided by infrastructures with medium to high complexity. It is designed for Service Level Monitoring, for medium and large infrastructures. It supports flexible development models and modular design, allowing easy communication with external systems (such as CMDBs, Service Catalogs, etc.). The service allows the creation of multiple reports using profiles (profiles are different ways and filters of organizing and displaying the services and their service components), which are defined by the client (e.g. for managing SLAs, operations, etc.). When calculating reports (using big data management technologies) the service considers a number of external factors, such as specific service parameters, scheduled or unscheduled downtime, etc.

The monitoring perspective of the integrated services is presented in Figure 1 through the monitoring dashboard.
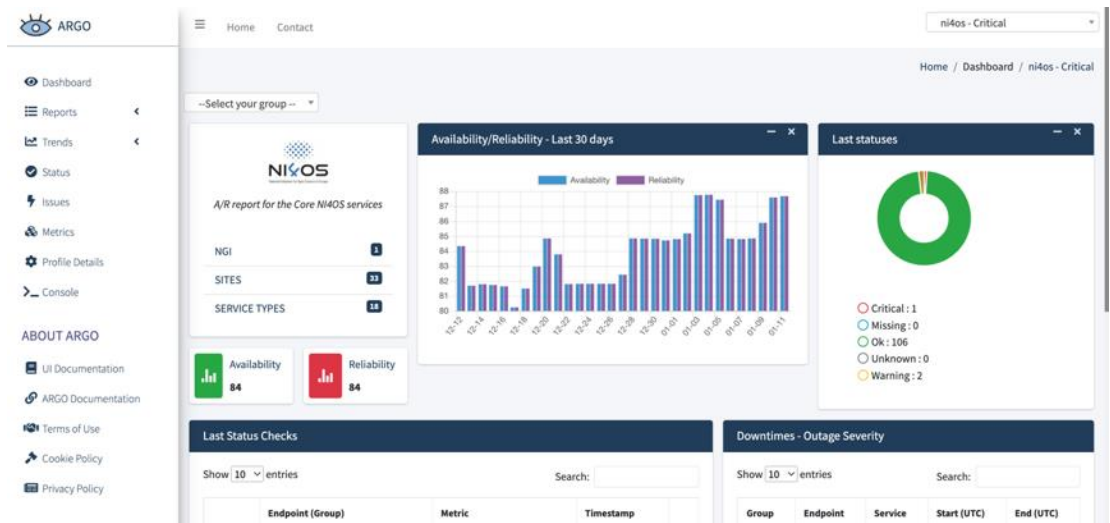
Figure 1: Dashboard of the Monitoring Service

The monitoring services also enables Availability/ Reliability report for the services and the Providers they belong, as shown in Figure 2.

Figure 2: Availability/ Reliability report for the services and the Providers they belong

It also displays the summary status of services per provider that provides them, enabling a wider overview of the status of each on-boarded providers, as shown in Figure 3.

Figure 3: Status report for the services and the Providers they belong

Another important element of the NI4OS-Europe pre-production environment is the accounting system. It enables the providers who feed usage data from their services to this system to get visual representations and follow the usage trends of their services.

webscience usage by Application and Date

12-2022	
■ REVIGO	7223
■ NanoCrystal	3862
■ ChemBioServer	616
■ FEPrepare	61
■ Ingredio	58
■ DREAMM	55
■ OMApp	21
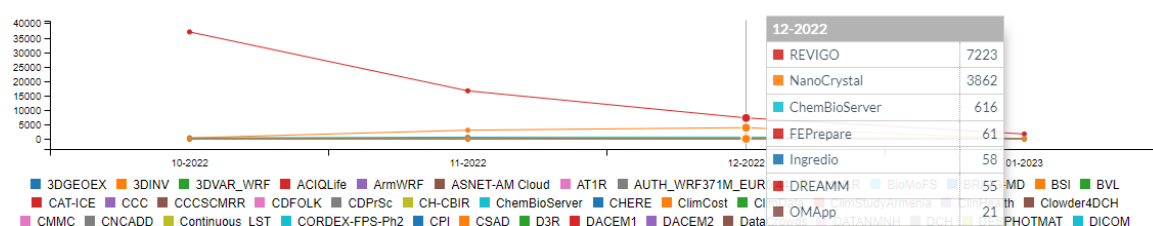
Figure 4: WebScience Application usage

The graphic in Figure 4 shows the usage of the thematic services(applications) that have integrated the accounting service. In this instance the graph shows the number of workflows for each application.
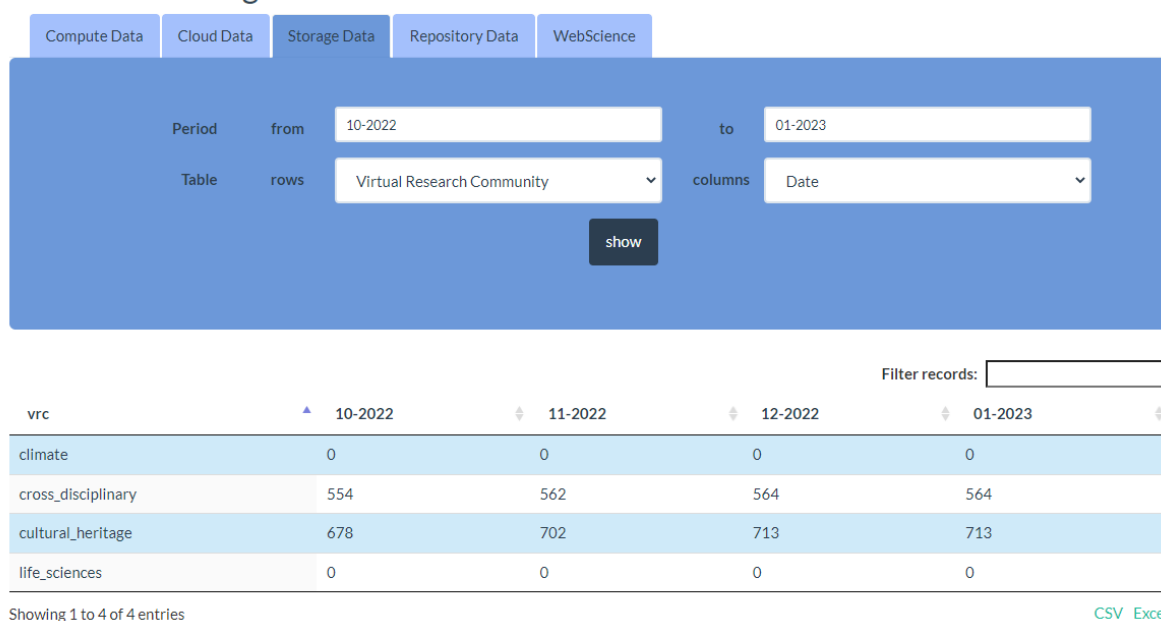
Figure 5: Storage usage by Research Community

In Figure 5 a table of the usage of storing services is shown. In this example the services are grouped by Research Community. The data displayed is in Gigabytes.

Total VM Hours by Resource name

26.6%

ETFBL-Cloud   26.6%

73.4%

■ Avitohol ■ Cyl Cloud ■ ETFBL-CC01 ■ ETFBL-Cloud ■ IIAP Cloud ■ InfinityCloud ■ InfraGRID Cloud ■ MD-RSC ■ MK-04-FINKI_CLOUD ■ Okeanos ■ UPT-Cloud
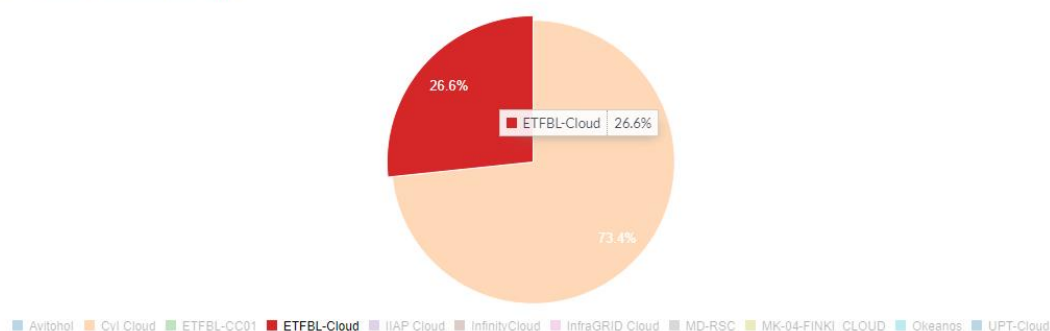
Figure 6: Cloud usage by hours

Figure 6 shows a pie chart representing the usage of cloud services by combining the total uptime of all virtual machines for each service.

One of the main aims of the project's helpdesk is to assess provided support. The helpdesk is fully integrated with the NI4OS-Europe on-boarding dashboard [7], and it is primarily utilized via this channel. From the dashboard, tickets are also accepted from external (non-registered) users by email, which significantly simplifies ticket creation and communication between support staff and the users. Up to this point, the total number of tickets created within the helpdesk reached 1,000. The time distribution of the created/closed tickets is illustrated in Figure 7. The average number of created and closed tickets per month is similar (38 tickets), so we can conclude that the response time of support units is high and that problems are resolved within a month, but practically much faster (average resolve time is 7 days). In the Figure 7, we can see three areas in which the number of created/closed tickets is much larger. The first area (beginning of the year 2021) corresponds to the initial service evaluation by user communities and the first set of providers and repositories on-boarding. From September 2021, we organized the on-boarding of the second set of providers and repositories, and these support activities produced the second area in Figure. Finally, the third peak in July 2022 corresponds to open call projects' support activities.
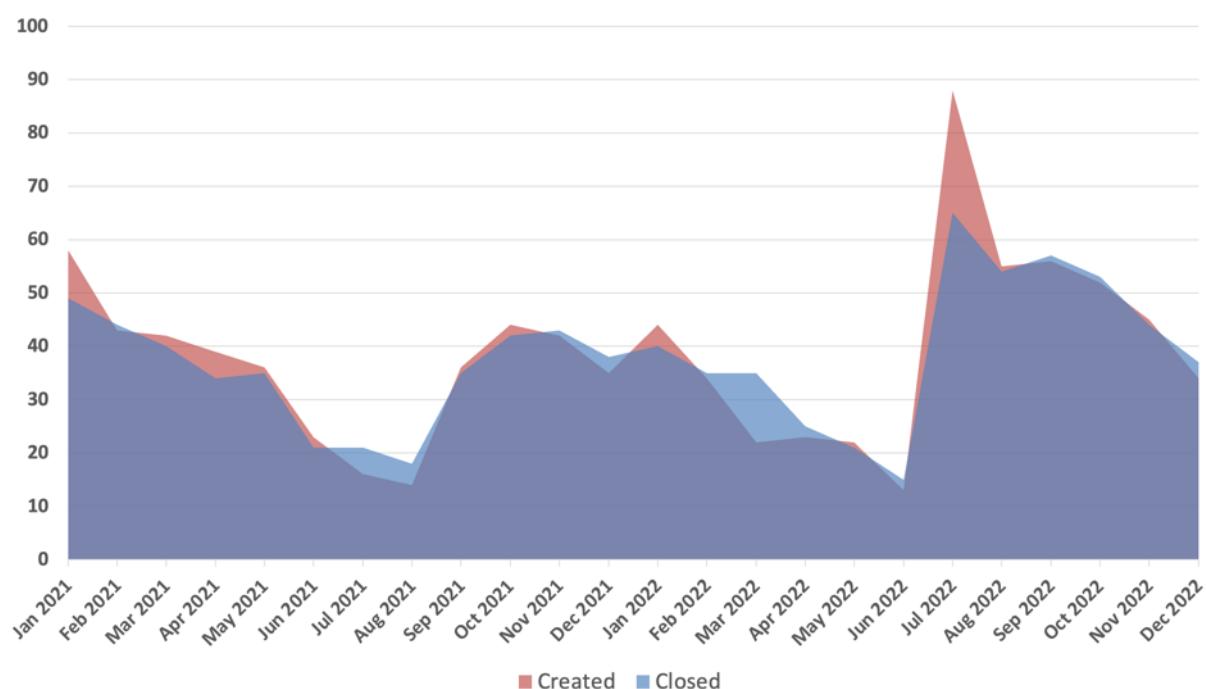
Figure 7: The number of created/closed tickets during the project's lifetime

During the summer of 2022, the NI4OS-Europe team migrated from the EOSC profiles specification from version 3 to version 4, which was a new development in the EOSC environment. Practically, this required an upgrade of the underlying schema used to describe registered services and repositories. The AGORA team finalized this migration, as well as the upgraded of the API and related documentation. There are several changes to the EOSC resource providers profile, which include: ADD new field "abbreviation" to the resource profile; ADD new field "multimedia name" to Providers/Resources; ADD "Use case name" new field to resources; MOVE hosting legal entity to basic info section; MOVE structure type to classifications section; UPDATE Hosting legal entity to show a dropdown list; UPDATE Related platforms to show a dropdown list; UPDATE multimedia fields to accept multiple entries; UPDATE Terms of Use and privacy policy to be mandatory fields for pushing to EOSC; UPDATE Service level field name to Resource level. On the side of the EOSC provider profile, changes from v3 to v4 are mostly reflected as movements of attributes from one to another block, as well as extension of the available values for the enumerated fields.

2. Updates of the on-boarding procedure

The main steps of the on-boarding process we defined at the project's beginning did not change much compared with the current procedure. Practically the skeleton of the procedure: information gathering, integration, validation, and publication, remained the same but was significantly fine-tuned and supported with additional operational tools. These main on-boarding steps are illustrated in Figure 8. In the first row, we listed on-boarding actions for which providers are responsible, and in the second row, actions performed with the support of the NI4OS-Europe on-boarding team. In the horizontal direction, on-boarding procedure steps are chronologically listed.
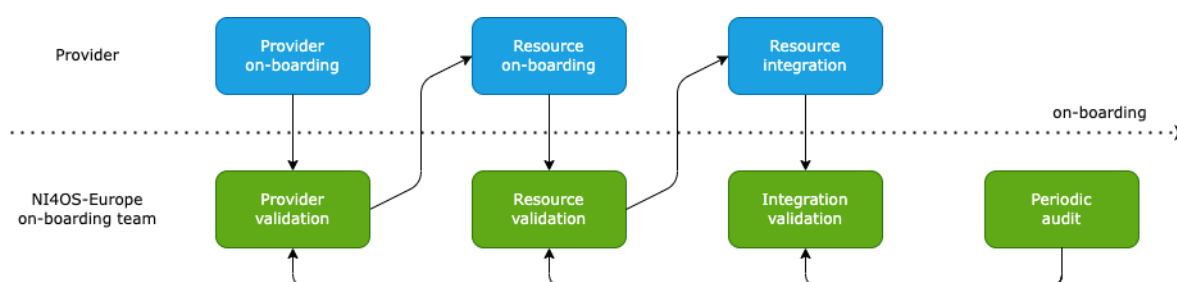
Figure 8: Main steps of the on-boarding procedure

The goal of the on-boarding process is to incorporate a resource outside of the EOSC into the EOSC federation, and once it is done, the resource will appear in the EOSC exchange layer [11]. Although the goal is the EOSC exchange layer, practical on-boarding includes unification with the tools from the EOSC core layer and integration with services from the exchange layer following guidelines from the EOSC interoperability layer. In this process, providers are supported by training material from the EOSC support layer. Therefore, within the NI4OS-Europe we have defined on-boarding as all practical activities to be taken to incorporate a research resource into the EOSC federation. These activities represent a wide range of support actions to be directly offered to the resource provider in the on-boarding process. For example, for a particular resource, this could be the establishment of the support channel, integration with the existing EOSC services, data FAIRification, integration with monitoring, accounting, or authentication/authorization frameworks, preparation of end-user tutorials, access policies, or terms of use, etc.

The main tool of the NI4OS-Europe on-boarding procedure is the AGORA catalogue, and practical on-boarding starts with this tool. AGORA is fully integrated with our Login service, so anyone can access the system using either institutional credentials or social identities. Once logged in, a customer will get default observer credentials and will be able to read registered providers' and resources' descriptions, which will give an image of how service descriptions should be prepared and lay out a learn-by-example approach. In order to be able to change or add new resource providers, the AGORA admin has to upgrade the default role to provider admin role. Similarly, in order to edit resource information, a customer will need resource admin credentials. Within the NI4OS-Europe, typically, we have one provider admin account per each provider and at least one AGORA resource admin account per registered resource. In addition, AGORA recognizes the Portfolio admin role, and customers with these credentials can validate resource and provider profiles and

then approve and publish them into the EOSC marketplace. The idea behind this procedure is to establish at least one portfolio manager role per country and in that way, to create a fully decentralized on-boarding procedure.

Once AGORA's account is upgraded from observer to admin role, the customer should proceed with the provider on-boarding first. Within the AGORA's provider section, by selecting a create button, the customer will be redirected to the corresponding form where provider details can be stored. In the form we expect a brief description of the provider, acronym, geographical location, website, domain, subdomain, set of tags that will be used by the search mechanism, and contact information. Once it is entered into the system, the information is verified by our on-boarding team, and the provider's status is changed from registered to on-boarded.

The next step of the on-boarding procedure is resource registration, which can be similarly completed using the AGORA's corresponding form. In order to be fully compatible with EOSC, we are using the latest EOSC profiles' specification, which is incorporated into our resource catalog. Practically, providers' and resources' forms used in the process of on-boarding follow this specification that ensures compatibility of descriptions. Similarly, to establish compatibility with the EOSC core layer, we deployed a set of compatible operational services to be able to measure the quality of the registered resources. In particular, we deployed AAI (or Login service), GOCDB, Helpdesk, Accounting, monitoring (ARGO), and training service, and our resource on-boarding procedure requires the integration with services from the EOSC core compatible layer, which is within the NI4OS-Europe called pre-production environment. This step is finalized by validation of the achieved integration, and once it is validated, the resource is considered on-boarded. However, to ensure that provided information is up-to-date and valid, all descriptions stored within the catalogue are periodically audited by the NI4OS-Europe on-boarding team.

3. Feedback on the usability of the pre-production environment

In order to collect continues feedback about project developments, a short survey for the on-boarded service providers was performed regarding the usability of the pre-production environment. The providers were asked to list the services from the pre-production environment that they used/integrated when on-boarding their services, as well as identify any problems and recommend changes to the environment.

As expected, the AGORA service portfolio management tool was the most frequently used tool from the pre-production environment, implementing the first step of the on-boarding procedure.

The monitoring tools was also used frequently by the providers to check the availability of their services, especially since for most of the services, no special effort was required from their side.

The Helpdesk services is crucial for providing the communication between the end-users and the service providers. This is especially visible in the previous paragraph for the period of the open call, where lots of tickets were opened from the users to acquire support from the service providers.

Accounting tool, with its latest extensions is covering more services types, offering the service providers more insight into the usage of their services. Through automation, the process of uploading the usage data is simplified, allowing better insight into the usage of the services.

Training and webinar system were also identified as important parts of the pre-production environment, providing platforms for the end-users to get familiar with the capabilities and usage scenarios of the services.

Regarding the recommendations, most of them were on the simplification and better explanation of the on-boarding procedures, as well as improving the cross links between the tools from the pre-production environment. All the recommendations were considered in detail and addressed by the hosting partners of the respective tools.

4. On-boarded providers

Through the project lifetime, 58 service providers from the countries participating to the project were successfully on-boarded to the NI4OS-Europe catalogue.

Figure 9 gives a snapshot of the interactive map of NI4OS-Europe providers in the SEE region. The map is available as part of the NI4OS-Europe on-boarding dashboard.

The current list of providers is presented in Table 1.
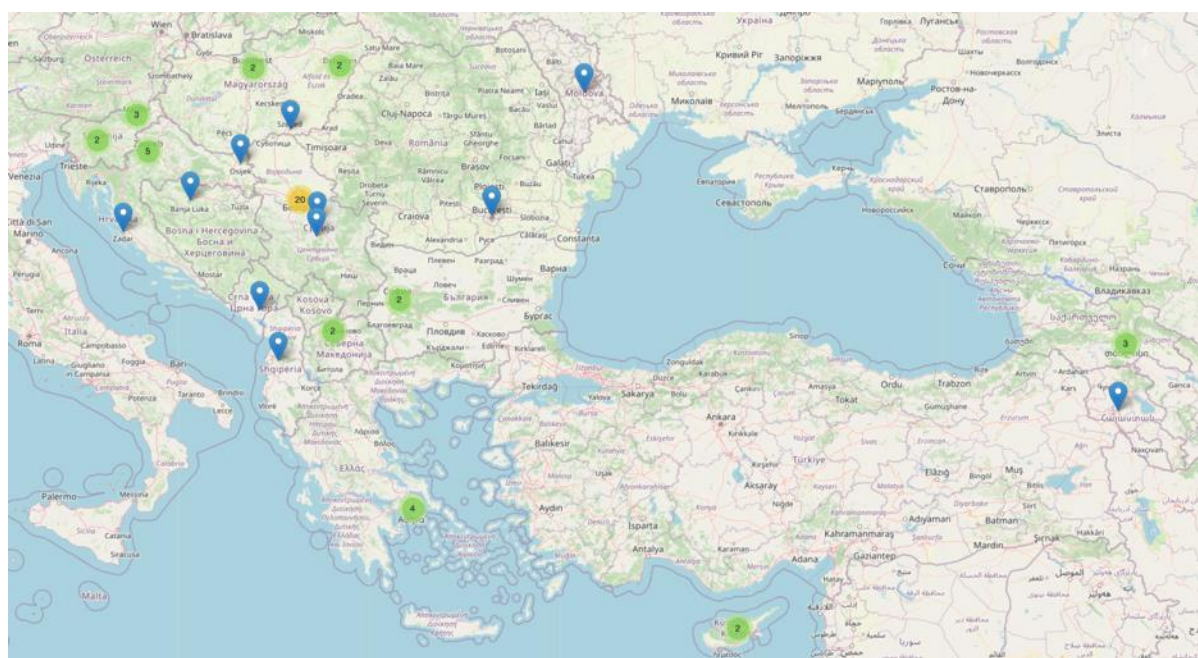
Figure 9: Map of NI4OS-Europe providers

Table 1: List of NI4OS-Europe service providers

Abbreviation	Country	Name
ARNES	SI	Academic and Research Network of Slovenia
RASH	AL	Albanian Academic Network - Interinstitutional ICT Research Centre
ATHENA	GR	ATHINA - Erevnitiko Kentro Kainotomias stis Technologies tis Pliroforias
BRFAA	GR	Biomedical Research Foundation, Academy of Athens
DHMZ	HR	Croatian Meteorological and Hydrological Service
PMF	HR	Faculty of Science, University of Zagreb
GRENA	GE	Georgian Research and Educational Networking Association
KIFU	HU	Governmental Agency for IT Development
IAH	RS	Institute for Animal Husbandry, Belgrade-Zemun
IBISS	RS	Institute for Biological Research Siniša Stanković, University of Belgrade
IIAP	AR	Institute for Informatics and Automation Problems
IFDT	RS	Institute for Philosophy and Social Theory, University of Belgrade

IPPE	RS	Institute for Plant Protection and Environment
IVCSP	RS	Institute for Vegetable Crops, Smederevska Palanka
IAUS	RS	Institute of Architecture and Urban and Spatial Planning of Serbia
IChTM	RS	Institute of Chemistry, Technology and Metallurgy, University of Belgrade
IICT-BAS	BG	Institute of Information and Communication Technologies, Bulgarian Academy of Sciences
IZUM	SI	Institute of Information Science
IMI-BAS	BG	Institute of Mathematics and Informatics at the Bulgarian Academy of Science
IPB	RS	Institute of Physics Belgrade
IJF	HR	Institute of Public Finance, Zagreb
IBCEB	GE	Ivane Beritashvili Center of Experimental Biomedicine
UNIOS EFOS	HG	J.J. Strossmayer University of Osijek, Faculty of Economics in Osijek
JSI	HG	Jožef Stefan Institute
MRIZP	RS	Maize Research Institute Zemun Polje
MESTD	RS	Ministry of Education, Science and Technological Development of the Republic of Serbia
GRNET	GR	National Infrastructures for Research and Technology
ICI	RO	National Institute for Research and Development in Informatics
NOA	GR	National Observatory of Athens
NSL-GE	GE	National Science Library at Tbilisi State University
OBM	HU	OpenBioMaps Consortium
RENAM	MD	Research and Educational Networking Association of Moldova
MTA TK KDK	HU	Research Documentation Centre, Centre for Social Sciences of the Hungarian Academy of Sciences
RBI	HR	Ruđer Bošković Institute
SANU	RS	Serbian Academy of Sciences and Arts
CyI	CY	The Cyprus Institute
CING	CY	The Cyprus Institute of Neurology and Genetics
UGD	MK	University Goce Delcev in Shtip
UNI BL	BH	University of Banja Luka
FAUB	RS	University of Belgrade - Faculty of Architecture
FCUB	RS	University of Belgrade - Faculty of Chemistry
FCE	RS	University of Belgrade - Faculty of Civil Engineering
STOMF	RS	University of Belgrade - Faculty of Dental Medicine
FPhB	RS	University of Belgrade - Faculty of Pharmacy
FSSUB	RS	University of Belgrade - Faculty of Security Studies
FVM	RS	University of Belgrade - Faculty of Veterinary Medicine
UoB-RCUB	RS	University of Belgrade Computer Centre
KPU	RS	University of Criminal Investigation and Police Studies

UD	HU	University of Debrecen
UNIKG	RS	University of Kragujevac
UM	SI	University of Maribor
UMUKM	SI	University of Maribor, University of Maribor Library
UoM	MN	University of Montenegro
SZTE	HU	University of Szeged
FSHN	AL	University of Tirana - Faculty of Natural Science
UNIZD	HR	University of Zadar
SRCE	HR	University of Zagreb, University Computing Centre (SRCE)
UKIM, FCSE	MK	University Ss. Cyril and Methodius, Faculty of Computer Science and Engineering
UKIM, FNSM	MK	University Ss. Cyril and Methodius, Faculty of Natural Science and Mathematics
INS Vinca	RS	Vinča Institute of Nuclear Sciences, Univeristy of Belgrade

5. On-boarded resources

NI4OS-Europe on-boarded resources are categorized into 4 main categories: repositories, thematic services, generic services and services from the pre-production environment. Following the on-boarding procedure described previously, 39 repositories, 33 thematic services, 23 generic services and 8 core or services from the pre-production environment have been successfully on-boarded.

5.1. On-boarded repositories

A total of 39 repositories from the countries participating in the NI4OS-Europe project were successfully on-boarded to the catalogue, following the latest EOSC provider profile metadata specification.

5.1.1.   NI4OS-Europe repository service

NI4OS-Europe 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.
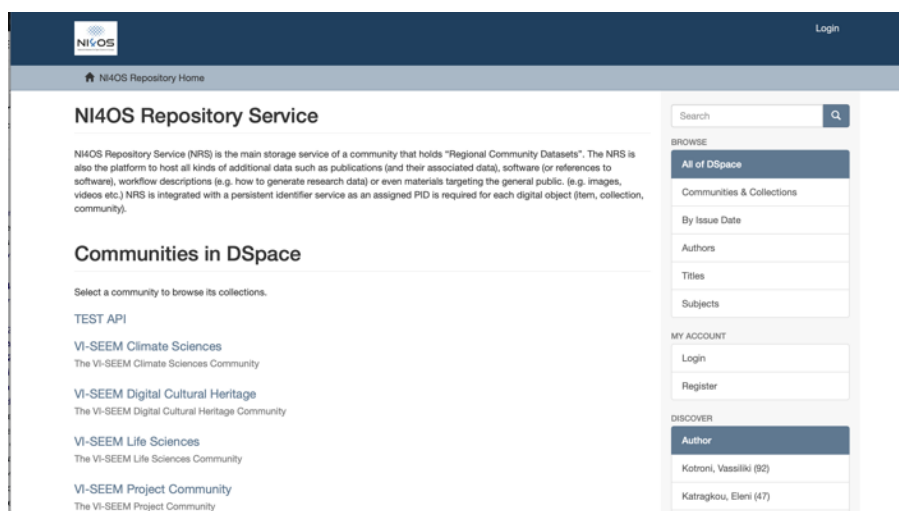
**Figure 10: NI4OS-Europe repository**

5.1.2.   CHERRY - CHEmistry RepositoRY

CHERRY, ie CHEmistry RepositoRY is a joint digital repository of the all departments in University of Belgrade - Faculty of Chemistry. CHERRY provides open access to the

publications, as well as to other outputs of the research projects implemented in this institution.

5.1.3.   CeR - Central Repository of the Institute of Chemistry, Technology and Metallurgy

CeR – Central Repository is the institutional digital repository of the Institute of Chemistry, Technology and Metallurgy, University of Belgrade. The aim of the repository is to provide open access to publications and other research outputs resulting from the projects implemented by the Institute of Chemistry, Technology and Metallurgy.

5.1.4.   DAIS - Digital Archive of the Serbian Academy of Sciences and Arts

DAIS - Digitalni arhiv izdanja SANU / Digital Archive of the Serbian Academy of Sciences and Arts is a joint digital repository of the Serbian Academy of Sciences and Arts (SASA) and the research institutes under the auspices of SASA. The aim of the repository is to provide open access to publications and other research outputs resulting from the projects and other activities implemented by the SASA and its institutes, as well as to ensure their long-term preservation. DAIS responds to the requirements of national and international funding bodies to share the outputs of publicly funded research and is compliant to the Open Science Platform of the Ministry of Education, Science and Technological Development of the Republic of Serbia.

5.1.5.   Digital Library of University of Maribor

The University of Maribor Digital Library (DKUM) is the institutional repository of the University of Maribor. It supports open access to scientific, research and professional works, and research data, which are results of research and education at the University. Publishing works in DKUM encourages social responsibility, exchange of scientific information, it impacts the effect of researchers' results, and the University's reputation. DKUM also enables the fulfilment of financiers' demands on open access to all reviewed publications and research data that are funded by public finances. The European Union prepared a Framework programme for funding the research activities and innovations, called Horizon 2020, which determines that all reviewed publications from funded project have to be published in open access. DKUM includes works from all University of Maribor members. Next to diplomas, master's degrees, doctorates and other works by students it also includes reviewed publications from funded projects, electronic academic textbooks and materials, and other works whose authors are the University of Maribor staff or if they are published by the University of Maribor.

5.1.6.   Digital Repository of Georgian Scientific Works

Georgian Integrated Library Information System Consortium GILISC was founded in 2017 to assist educational institutions in their development. GILISC gives them access to the electronic databases of scientific journals, helps them to use modern electronic catalogs/integrated library systems, store their scientific works in the digital repository (openscience.ge), publishes scientific articles of their students and staff in electronic journals (openjournals.ge), helps them to increase the quality of their student's scientific works. GILISC has been cooperating with the National Science Library of Georgia since its establishment and one of the results of their interaction is the repository openscience.ge.

5.1.7.   HELIX Data

HELIX Data is one of the services provided by the Hellenic Data Service. Hellenic Data Service, also known as HELIX, is a data catalogue and repository supporting knowledge management and scholarly communication in Greece. It is comprised of a diverse set of services, features and functionalities that facilitates responsible research conduct while easing compliance with Open and FAIR practices in the Greek research area. Moreover, HELIX succeeds in linking digital assets of publications, data and processes together thus contributing to information contextualisation and ultimately to building a European data economy of added value services and return on investments through improved research exploitation. HELIX content and tools are inclusive to all, from individual researchers and research communities to SMEs, data and citizen scientists and data enthusiasts. Their use is unlimited, as long as the licenses of the given/selected artefacts permit it, for research as well as educational or experimental purposes. HELIX Data was developed according to the principle of interdisciplinarity to serve the (complex) research needs of cross-disciplinary groups. For that, it supports a wide range of formats and standards and applies open and interoperable solutions for the sharing and preservation of its content in the long-term. Most importantly, HELIX encourages reuse and, since it is still a work in progress, at a later stage it provisions training activities around best practices on open and FAIR publishing and RDM for its users.

5.1.8.   RADaR - Digital Repository of Archived Publications of the Institute for Biological Research

RADaR - Digital Repository of Archived Publications of the Institute for Biological Research "Siniša Stanković" is the institutional digital repository of the Institute for Biological Research "Siniša Stanković" – National Institute of Republic of Serbia, University of Belgrade. The aim of the repository is to provide open access to publications and other research outputs resulting from the projects implemented by the Institute for Biological Research "Siniša Stanković".

5.1.9.   RIVeC - Repository of the Institute for Vegetable Crops

RIVeC - Repository of the Institute for Vegetable Crops Smederevska Palanka is the institutional digital repository of the Institute for Vegetable Crops. The aim of the repository is to provide open access to publications and other research outputs resulting from the projects implemented by the Institute for Vegetable Crops. The repository uses a DSpace-based software platform developed and maintained by the Belgrade University Computer Centre (RCUB). The software platform is compliant with the OpenAIRE Guidelines for Literature Repositories v3.

5.1.10.   Repository of Faculty of Science, University of Zagreb

The repository provides access to publications and research data produced by the employees and the students of the faculty. Content of repository is open to all users for searching and downloading with clearly stated usage rights.

5.1.11. Repository of the Institute of Public Finance, Zagreb

Repository provides access to publications and research data produced by the employees of the institute, papers published in scientific journals, conference proceedings, dissertations, books, manuals, guides and complete documentation related to the Institute's activities. Content of repository is open to all users for searching and downloading with clearly stated usage rights.

5.1.12. SZTE repository of publications

The intention of the SZTE Repository of Publications 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 viewability 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).

5.1.13. Social Scientific Research Documentation Centre Repository

The Research Documentation Centre of the Centre for Social Sciences provides information on and access to research conducted at the Centre. The metadata and some of the documents of the Research Documentation Centre (RDC) are available to all visitors, but many are restricted to registered users. The interface can be set to many languages and contains RSS feeds to alert users of new content.

5.1.14. University of Zadar Institutional Repository of evaluation works

The repository provides access to publications and research data produced by the employees and the students of the faculty. Content of repository is open to all users for searching and downloading with clearly stated usage rights.

5.1.15. VideoLectures.Net

VideoLectures.NET is an award-winning free and open access educational video lectures repository. The lectures are given by distinguished scholars and scientists at the most important and prominent events like conferences, summer schools, workshops and science promotional events from many fields of Science. The portal is aimed at promoting science, exchanging ideas and fostering knowledge sharing by providing high quality didactic contents not only to the scientific community but also to the general public. All lectures, accompanying documents, information and links are systematically selected and classified through the editorial process taking into account also users' comments.

5.1.16. VinaR - Repository of the Vinča Institute of Nuclear Sciences

VinaR - Repository of the Vinča Institute of Nuclear Sciences is the institutional digital repository of the Vinča Institute of Nuclear Sciences, University of Belgrade. The aim of the repository is to provide open access to publications and other research outputs resulting from the projects implemented by the Vinča Institute of Nuclear Sciences.

5.1.17. FarFar - Repository of the Faculty of Pharmacy

FarFar - Repository of the Faculty of Pharmacy is the joint digital repository of all departments of the University of Belgrade, Faculty of Pharmacy. The aim of the repository is to provide open access to publications and other research outputs resulting from the projects implemented by the Faculty of Pharmacy and its departments.

5.1.18. GraFar - Repository of the Faculty of Civil Engineering

GraFar - Repository of the Faculty of Civil Engineering is the joint digital repository of all departments of the University of Belgrade - Faculty of Civil Engineering. The aim of the repository is to provide open access to publications and other research outputs resulting from the projects implemented by the Faculty of Civil Engineering and its departments.

5.1.19. Jakov - Repository of the University of Criminal Investigation and Police Studies

Jakov - Repository of the University of Criminal Investigation and Police Studies in Belgrade is the institutional digital repository of the University of Criminal Investigation and Police Studies in Belgrade. The aim of the repository is to provide open access to publications and other research outputs resulting from the projects implemented by the University of Criminal Investigation and Police Studies.

5.1.20. Meteorological and Hydrological Service of Croatia Repository

The repository provides access to publications and research data produced by the employees of the institute. Content of repository is open to all users for searching and downloading with clearly stated usage rights.

5.1.21. NaRDuS - National Repository of Dissertations in Serbia

NaRDuS (National Repository of Dissertations in Serbia) is a common portal of PhD dissertations and thesis evaluation reports from all Serbian universities. It is based on the Law on Higher Education (Amendments, Sept. 2014).All universities are obliged to deposit basic information about the dissertation - together with the thesis evaluation report and dissertation itself - to NaRDuS within three months period starting from the date of PhD dissertation defense.

5.1.22. National Bibliometric Instrument

National Bibliometric Instrument (IBN) At the moment, IBN is the biggest Open Access electronic library of articles published in national scientific journals, as well as conference proceedings from the Republic of Moldova starting from 1993 to date. From 2022, IBN includes data on monographs, book chapters and sections. The interface is available in Romanian.

5.1.23. PlantaRum – Repository of the Institute for Plant Protection and Environment

PlantaRum – Repository of the Institute for Plant Protection and Environment is the institutional digital repository of the Institute for Plant Protection and Environment in

Belgrade. The aim of the repository is to provide open access to publications and other research outputs resulting from the projects implemented by the Plant Protection and Environment.

5.1.24.   RAF - Repository of the Faculty of Architecture

RAF - Repository of the Faculty of Architecture is the joint digital repository of all departments of the University of Belgrade, Faculty of Architecture. The aim of the repository is to provide open access to publications and other research outputs resulting from the projects implemented by the Faculty of Architecture and its departments.

5.1.25.   RAUmPlan - Repository of Architecture, Urbanism and Planning

RAUmPlan - Repository of Architecture, Urbanism and Planning is the institutional digital repository of the Institute of Architecture and Urban & Spatial Planning of Serbia. The aim of the repository is to provide open access to publications and other research outputs resulting from the projects implemented by the Institute of Architecture and Urban & Spatial Planning of Serbia.

5.1.26.   RIFDT - Repository of the Institute for Philosophy and Social Theory

RIFDT - Repository of the Institute for Philosophy and Social Theory is the institutional digital repository of the Institute for Philosophy and Social Theory of the University of Belgrade. The aim of the repository is to provide open access to publications and other research outputs resulting from the projects implemented by the Institute for Philosophy and Social Theory.

5.1.27.   RIK - Repository of the Maize Research Institute Zemun Polje

RIK - Repository of the Maize Research Institute Zemun Polje is the digital repository of the Maize Research Institute Zemun Polje. The aim of the repository is to provide open access to publications and other research outputs resulting from the projects implemented by the Maize Research Institute Zemun Polje.

5.1.28.   RIStocar - Repository of the Institute of Animal Husbandry

RIStocar – Repository of the Institute of Animal Husbandry is the institutional digital repository of the Institute of Animal Husbandry in Belgrade. The aim of the repository is to provide open access to publications and other research outputs resulting from the projects implemented by the Institute of Animal Husbandry.

5.1.29.   Repository of the Faculty of Economics in Osijek

The digital repository gathers, permanently stores and allows open access to digital versions of scientific research, intellectual and creative property of an institution, its employees and students. The repository stores students' M.A. and B.A. theses, dissertations, pre-print papers, scientific and professional papers, research data, books, teaching materials, images, video and audio files, presentations and digitized material.

5.1.30.   SCIDAR - A Digital Archive of the University of Kragujevac

A digital archive of the University of Kragujevac contains easily searchable and up-to-date research results. The institutional digital repository of the University of Kragujevac was set up in February 2020. An initial collection of primary scientific results is available.

5.1.31.   SMILE – School of dental Medicine dIgitaL archive

SMILE – School of dental Medicine dIgitaL archivE is the institutional digital repository of the University of Belgrade, School of Dental Medicine. The aim of the repository is to provide open access to publications and other research outputs resulting from the projects implemented by the School of Dental Medicine.

5.1.32.   Veterinar – Repository of the Faculty of Veterinary Medicine

Veterinar – Repository of the Faculty of Veterinary Medicine is the institutional digital repository of the University of Belgrade, Faculty of Veterinary Medicine. The aim of the repository is to provide open access to publications and other research outputs resulting from the projects implemented by the Faculty of Veterinary Medicine.

5.1.33.   Albanian Research Publications Repository

The U-CRIS platform is the database of Albanian universities for storing scientific publications, doctorates, CVs of research staff, research projects, etc. The platform is based on the DSPACE-CRIS architecture which has been adapted according to EUROCRIS and CERIF specifications. The platform is linked to important search databases such as PubMed, Scopus, Web of Knowledge (WoS), Google Scholar and ORCID. U-CRIS also enables connection to European open Aire search databases. The U-CRIS platform will also be linked to the ACRIS national platform envisaged for the classification of HEIs in the field of scientific research.

5.1.34.   Armenian NOAD

The scientific repository of the Institute for Informatics and Automation Problems of the National Academy of Sciences of Armenia providing access to the research publications of the institute. The interface is available both in Armenian and English.

5.1.35.   Institutional Repository in Medical Sciences

Institutional Repository in Medical Sciences of Nicolae Testemitanu State University of Medicine and Pharmacy of the Republic of Moldova (IRMS – SUMPh) collects, preserves and gives open access to conference papers, articles from scientific journals, reports, theses and dissertations, study aids, patents, student theses, and other publications created by members of the Nicolae Testemitanu University of Medicine and Pharmacy.

5.1.36.   Institutional Repository of Economic Knowledge

Institutional Repository of Economic Knowledge provides access to the research output of the Academy of Economic Studies of Moldova. The interface is available in Romanian.

5.1.37.   RhinoSec - Repository of the Faculty of Security Studies

RhinoSec - Repository of the Faculty of Security Studies is the digital repository of the the University of Belgrade, Faculty of Security Studies. The software platform of the repository is adapted to the modern standards applied in the dissemination of scientific publications and is compatible with international infrastructure in this field.

5.1.38.   Ruđer Bošković Institute Research Data Repository

FULIR Data is a research data repository that gathers, permanently stores and allows open access to primary data produced by researchers based at Ruđer Bošković Institute. Researchers deposit datasets by themselves (self-archiving) with the support given by the Centre for Scientific Information and their RDM experts.

5.1.39.   UKIM Repository

The repository of the University of Ss Cyril and Methodius in Skopje, containing scientific publication and scientific data (in progress), as well as the CRIS database for the University.

5.2. On-boarded thematic services

On-boarded thematic services within the NI4OS-Europe catalogue cover many different scientific domains, with major being life sciences, climate science, digital cultural heritage and computational physics.

5.2.1.   Life sciences thematic services

The live sciences thematic services category consists of 18 services providing rich set of possibilities to the researchers of this scientific field.

5.2.1.1 ChemBioServer 2.0

ChemBioServer 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.
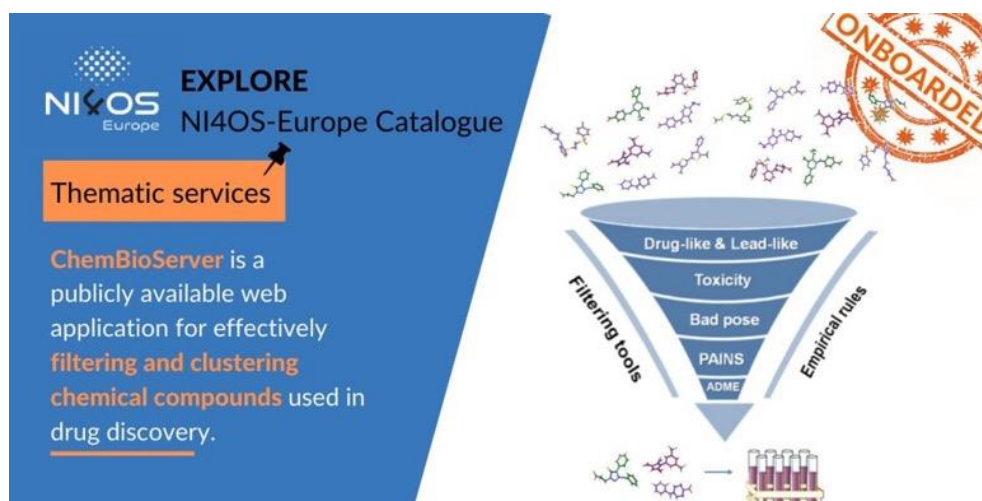
Figure 11: ChemBioServer

5.2.1.2 *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.

5.2.1.3 *DREAMM*

DREAMM 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.

5.2.1.4 *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.

5.2.1.5 *Ingredio*

Ingredio application is a natural processing language (NLP) application that offers a pipeline of three services related to biomedical text. The application is able to classify biomedical text based on certain features of its content, extract compound names and infer causal relations from the text, however it is experimental and is not meant to replace

human curation. It's main use is to showcase how this can be used as a high-throughput and high precision language filtering software for large scale biomedical data.

5.2.1.6 *OpenBioMaps*

OpenBioMaps is a web-based, open-access database framework project which maintained by the OpenBioMpas Consortium and the databases involved at least partially open-access or contains free-content. The OpenBioMaps provides an open-access web application which designed to create and use open-content biological databases, specifically for scientists and conservationists, and its customizable toolset allows for the easy access and management of data.

5.2.1.7 *REVIGO*

REVIGO is a web server that summarizes long, unintelligible lists of Gene Ontology (GO) terms by finding a representative subset of the terms using a simple clustering algorithm that relies on semantic similarity measures. Furthermore, REVIGO visualizes this non-redundant GO term set in multiple ways to assist in interpretation: multidimensional scaling and graph-based visualizations accurately render the subdivisions and the semantic relationships in the data, while treemaps and tag clouds are also offered as alternative views.

5.2.1.8 *EEGHUB*

EEGHUB.GE (http://eeghub.ge) is a Big Data EEG (Electroencephalogram-Brain Electrical Activity) online dataset in Georgia. The service is free for European (or national) researchers following the principles of findability and accessibility. Service has convenient search engine, which allows users to identify any recordings that correspond to specific requirements. The recordings are easily accessible and can be downloaded for further exploitation. The target users are open-source groups of researchers/practitioners, lecturer/students, Scientific Organization, Hospitals, Universities, etc. It is envisaged that EEG collection "eeg.hub.ge" will support researchers in the field of neuroscience, psychophysiology, medicine, psychology, neurophysiology, cognitive and social science.

5.2.1.9 *DICOM*

"DICOM Network" service provide full set of functionalities for data collect, storage, distribution and exchange DICOM medical investigations. It's implements all the standard PACS interfaces as well as integrated security features. various imagistic investigations like tomography, Roentgen, ultrasound, angiography, etc. Familiarization and working experience accumulation by medical specialists in using such systems offer obvious advantages in imagistic investigations and forming treatment decisions, allow supporting collaborative work and appealing for support from the best local and foreign specialists who have extensive experience in the field.

5.2.1.10   *MelGene*

The MelGene database provides a comprehensive and regularly updated field synopsis and meta-analysis of all published genetic association studies performed in Cutaneous Melanoma (CM). In addition, dozens of up-to-date meta-analyses are available for all eligible polymorphisms with sufficient data.

5.2.1.11   Computational Drug Repositioning Score

Drug repurposing techniques allow existing drugs to be tested against diseases outside their initial spectrum, resulting in reduced cost and eliminating the long time-frames of new drug development. In silico drug repurposing further speeds up the process either by proposing drugs suitable to invert the transcriptomic profile of a disease or by indicating drugs based on their common targets or structural similarity with other drugs with similar mode of action. Such methods usually return a number of potential repurposed drugs that need to be tested against the disease in in vitro, pre-clinical and clinical studies. Thus, it is crucial to have a more sophisticated candidate drug ranking in order to start testing from the most promising chemical substances. As a means to enhance the above decision process, we present CoDReS (Composite Drug Reranking Scoring), a drug (re-)ranking web-based tool, which combines an initial drug ranking (i.e. repurposing score or hypothesis/potentiality score) with a functional score of each drug considered in conjunction with the disease under study as well as with a structural score derived from potential drugability violations. Furthermore, a structural similarity clustering is applied on the considered drugs and a handful of structural exemplars are suggested for further in vitro and in vivo validation. The user is able to filter the results further, through structural similarity examination of the candidate drugs with drugs that have failed against the queried disease where related clinical trials have been carried out.

5.2.1.12   ProtExA

ProtExA is a web tool for protein post-processing analysis from proteomics experiments acting as an all-in-one tool for the statistical and functional analysis of protein expression datasets. It facilitates protein differential expression analysis and protein co-expression network analysis as well as pathway and post-pathway analysis through network-based bioinformatics approaches. ProtExA combines four main pillars of omics data analysis, commonly used in bioinformatics pipelines, providing significant information about the functional relations between protein and genes.

5.2.1.13   Virus to Drug Networks

Vir2Drug is a Drug Repurposing web tool that uses network-based approaches to identify and rank candidate drugs for a specific pathogen, combining information obtained from:

- ranked pathogen-to-pathogen (P2P) networks based on protein similarities between pathogens at host and/or proteome level,
- taxonomy distance between pathogens,
- drugs targeting specific host and/or proteome proteins.

Vir2Drug allows the creation of P2P similarity networks based either at host or proteome level, where the edges are scored by means of specific equations depending on whether the pathogen under study is categorised or not. P2P networks are used to screen drugs by means of six available methodologies that account for either the host or pathogen's protein targets, while the drug ranking is performed by means of seven proposed available equations. Vir2Drug comes with a frontend web interface that consists of the mainframe and a help page, written in HTML, PHP and JavaScript language environments. The mainframe provides 3 individual steps designed to guide the user until the end of the workflow process. The backend of Vir2Drug has been written in R environment, where several functionalities have been parallelised to achieve fast performance. Evaluation,

testing and understanding of Vir2Drug functionalities can be easily performed by means of several available examples provided on the web site.

5.2.1.14   *LOCKDOWN SCENARIOS TOOL based on ATVBG-SEIR Model*

How long "lockdowns" are needed to end the Covid-19 epidemic, with or without vaccinations. Web-based Lockdown Scenarios Tool. Results are demonstrating the efficiency of the tool by applying it to Covid-19 data from Austria, Bulgaria, Germany, Italy, UK, and USA.

5.2.1.15   *PATHWAY-CONNECTOR*

PathwayConnector is a web-tool that facilitates the construction of complementary pathway-to-pathway networks and subnetworks of them, based on a reference pathway network derived from the rich information available either in KEGG or Reactome database for pathway mapping. Specifically, for a given set of pathways, PathwayConnector (i) finds all the direct connections between them, (ii) adds a minimum set of complementary pathways required to achieve connectivity between the pathways, leading to informative fully connected networks and (ii) provides a series of clustering methods for the further grouping of pathways in to sub-clusters. The proposed web-tool is a simple yet informative tool towards identifying connected groups of pathways that are significantly related to specific diseases.

5.2.1.16   *PathExNet*

PathExNET is a web service that allows the creation of pathway expression networks that hold the over- and under-expression information obtained from differential gene expression analyses. PathExNET holds a large database of reference pathway-to-pathway networks, which have been developed through the freely available information included in the KEGG, Reactome and Wiki Pathways database repositories. Users can upload their differential gene expression statistical analysis, followed with pathways and/or genes of interest, and further chose a score methodology to create and explore the derived pathway-to-pathway expression networks. In order to provide a concrete set of well-evaluated differential gene expression statistical analyses and to further increase the data-availability and easy data access of PathExNET, an additional tool has been rooted in PathExNET framework that allows to search and directly import, pre-processed statistic files from the Expression Atlas (EA) (https://www.ebi.ac.uk/gxa) data repository of the European Bioinformatics Institute (EMBL-EBI) (https://www.ebi.ac.uk).

5.2.1.17   *PathIN*

PathIN is a web-service that provides an easy and flexible way for rapidly creating pathway based networks, at several functional biological levels: genes, compounds and reactions. The proposed tool holds a large database repository of reference pathway networks, across a large set of species, which have been developed through the freely available information included in the KEGG, Reactome, and Wiki Pathways database repositories. PathIN provides networks by means of five diverse methodologies: (a) direct connections between pathways of interest, (b) direct connections as well as the first neighbours of the given pathways, (c) direct connections, the first neighbours and the connections in between them, and (d) two additional methodologies for creating complementary pathway-to-pathway networks that involve additional (missing) pathways that interfere in-between pathways of interest. The underlying tool is expected to be used as a simple

yet informative reference tool for understanding networks of molecular mechanisms related to specific diseases. PathIN requires four general steps to search for specific pathways of interest and create their network.

5.2.1.18   *Atlas of prokaryotic traits*

The ProTraits a web-based atlas of prokaryotic traits describes environmental preferences of microbes, interactions with other organisms (including pathogenicity), biochemical phenotypes, resistance to chemicals and other stressors, and utility in industrial applications. ProTraits recognizes 424 phenotypic traits and covers 3,046 bacterial or archaeal species. Overall, it provides 545,081 annotations (less than 10% FDR, tallying both the positive and the negative labels), of which 503,308 are novel.

5.2.2.   Climate sciences thematic services

A total of 9 climate science thematic services have been made available through the NI4OS-Europe catalogue.

5.2.2.1 *Live Access Server*

The LAS enables the data provider to unify access to multiple types of data in a single interface, create thematic data servers from distributed data sources, and offer derived products on the fly. In the initial window screen, the user can select a new dataset using the "Data Set" button on the top left of the webpage. Using the zoom and pan reference map, the user may refine the latitude and longitude used to create the image or download data. The output images are interactive, allowing the user to click and drag to zoom into a particular region on the map.
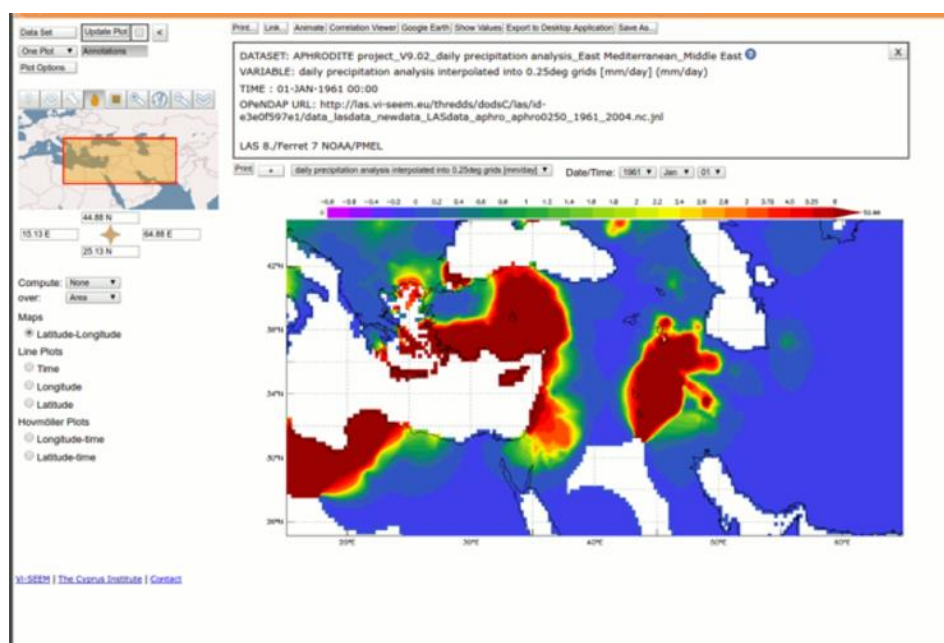

Figure 12: Live Access Server

5.2.2.2 OMApp

OMApp is a cloud application for automatic image mosaicking and georeferencing. The application is designed to support multiple users, where each user can upload a set of captured images through the web interface, start their processing, and preview created maps. Processed images can also be downloaded in standard open formats. OMApp uses a number of open-source image processing tools and libraries, where the most computationally demanding of them are able to perform multi-core parallel processing, allowing better use of cloud resources.

5.2.2.3 Remote Sensing Scene Clasification

RS2C is a RESTful web service and web application for remote sensing scene classification based on convolutional neural networks. Currently, ResNet-50 pre-trained on ImageNet and fine-tuned on MLRSNet is used for classification. The web service is implemented in Python using TensorFlow Serving and Flask. The RS2C API provides methods for single- and multi-label classification.

5.2.2.4 SciRoHub

The Service provides Copernicus's Sentinel1 and Sentinel2 products that cover the Romanian territory. Sentinel1 mission provides all-weather, day-and-night radar imaging for land and sea. Sentinel2 provides high resolution optical image data, including monitoring of vegetation, soil and water cover, as well as observation of inland waterways and coastal areas. It targets researchers from Romanian National Research Institutes and Universities that are interested in using satellite images for their research purposes. The Service is free to use for Romanian academic and research community. Registration is mandatory in order to access the data.

5.2.2.5 Air pollution Prediction

Simulation system for generation of prediction of air pollution levels based on WRF-Chem software. The outputs are hourly levels of airpollution for PM10, PM2.5, NO and SO2.

5.2.2.6 ClimCost

The service produces reliable, comprehensive and detailed evaluations of possible regional/local climate changes and their consequences for different global change scenarios. Metrics and tools for evaluating some of the climate change impacts on environment and quality of will be available. Making use of vast computing resources, the service will enable scientist to perform in-depth assessment of the climate change impacts that cannot be achieved with the desired accuracy using local computing resources. Users are climate scientists and national and municipal policymakers.

5.2.2.7 ClinHealth

The service generates reliable, comprehensive and detailed studies of the impact of lower atmosphere parameters and characteristics on the quality of life and health risks for the population in our country. It uses a synergetic application of extensive computer simulations on the supercomputer Avitohol, combined with sophisticate analysis of the parameters and characteristics of near surface atmosphere. In this way the impact of the atmosphere on the human health and quality of life can be thoroughly investigated. The

users are cross-disciplinary, from the domains of atmospheric physics and environmental science, allergology and epidemiology, as well as national and municipal policymakers.

5.2.2.8 DREAM

The DREAM service simulates and predicts the atmospheric cycle of mineral dust aerosols. It is tuned for usage on high-performance computing infrastructures available today. A typical use-case is the production of a dataset with the aerosol optical thickness and surface dust concentration for a particular period and for a particular geographical region. The service supports different horizontal and vertical resolutions. The results produced by the DREAM service have been applied, using the human health impact function and calculated global fine particulate matter concentrations, for estimation of the premature mortality caused by the long-term exposure to airborne desert dust. The results have high sensitivity on the threshold concentration, which is a significant parameter of relevance to public health.

5.2.2.9 High Energy Solar Particle Events foRecastIng and Analysis (HESPERIA)

Novel operational solar weather forecasting tools based upon proven concepts (UMASEP, REleASE). Advancement of understanding of the physical mechanisms that result into high-energy solar particle events (SEPs) exploiting novel datasets (FERMI/LAT/GBM; PAMELA; AMS) incorporation of the derived results into innovative space weather services. Customers include NASA and ESA.

5.2.3.   Digital Cultural Heritage services

In the field of Digital Cultural Heritage, two services from the region were on-boarded, with their detail shown below.

5.2.3.1 CHERE

CHERE Tools stands for Cultural Heritage Repository Tools and represents a set of web based tools aimed at people working in cultural heritage preservation and digitization, but is not limited to those uses as individual services can be used in a variety of ways. The service currently provides following functions:

- Structure from Motion - reconstruction of textured 3D object from a series of photographs of the object.
- Measurement of 3D objects - either reconstructed or user-provided.
- Conversion of BigBlueButton meetings' recordings to standalone video files.

5.2.3.2 Chowder4DCH

Clowder4DCH, a highly extensible active curation-based research data management platform. It contains three major extension points: preprocessing, processing and previewing. When new data is added to the system, preprocessing is off-loaded to extraction services for extracting appropriate data and metadata. The extraction services attempt to extract information and run preprocessing steps based on the type of the data, e.g., to create previews. This raw metadata is presented to the user via a web interface. Users can upload, download, search, visualize research datasets and explore information linked to data. Users can link and organise datasets in online collections following the provided workflows for creating semantically structured data repositories specialized for

digital cultural heritage. It enables users to form an online collaboration environment to support research communities and activities, and disseminate results.

5.2.4.   *Computational physics thematic services*

Two computational physics API based services have been on-boarded through the NI4OS-Europe service catalogue.

5.2.4.1 *Gaussian API*

Gaussian regression (GPR), along with other emerging machine learning techniques, has become more and more popular in computational chemistry, physics, biology and life sciences. In conjunction with the molecular dynamics simulations (MD), these approaches have been shown to be rather useful for prediction of a wide variety of molecular and materials' properties and functionalities. However, due to the novelty of techniques, the procedures for their application as well as their validation are far from being standardized.

5.2.4.2 *Schrödinger API*

n many subdisciplines of computational molecular sciences, computational physics, chemistry, biology, materials science, exact treatment and analysis of a wide variety of phenomena has to rely on rigorous quantum description of the underlying processes. Numerous phenomena taking place in the nano-world are inherently quantum in nature. Their description and, more important, quantitative treatment, therefore, requires usage of the apparatus of quantum mechanics. The basic paradigm of today's "mainstream" quantum mechanics is the Schrödinger equation, which is considered as a "quantum analog" to the famous Newton's second law equation in classical physics. The effort required to solve the Schrödinger equation is heavily dependent on the dimensionality and complexity of the problem itself (e.g. the exact form of the Hamiltonian, number of the relevant degrees of freedom of the studied system etc.). Numerous methods have been proposed in the literature to achieve the mentioned aim. However, the available codes are most often user-hostile, the procedures for computation and generation of relevant data are non-standardized, and there is a clear lack of in-depth, thorough comparison of performances of various methods for solving the Schrödinger equation for various purposes. The proposed service will provide user-friendly (as much as possible) computational platforms for solution of time-independent Schrödinger equation, implementing several algorithms that uses the Hermite discrete variable representation technique (DVR) approach.

5.2.5.   *Other disciplines thematic services*

In the group of other disciplines thematic services, there are 2 on-boarded services.

5.2.5.1 *SexEst*

SexEst is a free, interactive, web application designed to estimate sex using cranial or postcranial linear measurements. Users can either enter manually the measurements for single skeletons or upload data for multiple skeletons stored in a CSV file. Sex estimation is based on three different machine learning classification algorithms: Linear Discriminant Analysis (LDA), Extreme Gradient Boosting (XGB), and Light Gradient Boosting (LGB).

5.2.5.2 *Albanian Research Statistics Portal*

E-Research is the Portal where you can find statistical data about scientific research in HEIs, Research Institutes, Technology transfer Centers, Academy of Science, different Agencies under Ministries of lines in Albania. It summarizes statistics for:

- Human Research Capacities in HEIs
- Archives and scientific Libraries
- Research infrastructure and Laboratories

5.3. On-boarded generic services

A total of 23 generic services from the partnering countries in the NI4OS-Europe project were on-boarded to the service catalogue, providing computational and storage fabric to the researchers in the form of HPC, Cloud, Storage and analytics services.

5.3.1.   HPC generic services

5.3.1.1 *ARIS*

GRNET (National Infrastructures for Research and Technology) provides high performance computing resources to the Greek and international scientific and research communities in order to conduct scientific research. It is a typical HPC system. All compute nodes are used through SLURM resources/workload manager, they are not directly accessible by end user and they havn't internet access. All compute jobs run through SLURM. Access to the system is allowed only via SSH from specific Ips/networks to login nodes from which all data management/transfers, job submission etc. are performed. Only local LDAP authentication is supported. Authorization to use each partition, run limits, budget, accounting etc. are handled by SLURM. System has total 2 PB (raw) shared storage (gpfs) with usable, after raid etc. capacity of 1.4 PB splitted in three partitions with tuning for different types of storage usage. It supports large number of applications, optimized for each partition hardware. Software is organized via environment modules. It supports containers (ONLY singularity). Its average yearly usage is 86.6% during last year – including downtimes.

5.3.1.2 *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.

5.3.1.3 *Leo HPC*

IFÜ has been operating its supercomputer infrastructure since 2001. The components of this infrastructure (computational and data storage facilities) can be found in 5 locations. Leo, our most powerful HPC site, is a cluster type machine located in Debrecen, the second largest city in Hungary. The overall processing power of the 1344 CPU cores and the 252

GPUs is around 254 Tflop/s. Concerning the storage capacity, a total of 585 TB can be used for computational purposes. More than 380 projects have used the Leo cluster so far, and nearly 850,000 tasks have been submitted since its start. Approximately 70 disciplines have been registered over the years. Theoretical/computational chemistry, biology, and physics dominate resource allocations. Any person or research group whose host institution has a contractual relationship with KIFÜ can use the HPC facilities. Such institutions are typically universities, the Hungarian Academy of Sciences, museums, and special libraries. They can use the HPC infrastructure free of charge. The name Leo was chosen to pay homage to Leo Szilárd, a Hungarian physicist and inventor, although the name simultaneously refers to the lion, which symbolises strength and courage.

5.3.1.4 *Isabella*

Isabella provides Croatian researchers and their EU peers with access to high performance compute & storage resources. Isabella consists of 135 worker nodes with 3100 processor cores, 12 GPUs and 756 TiB data space. As a shared resource of all scientists in Croatia, it allows using significant computational resources in demanding data processing of scientific and research projects.

5.3.1.5 *PARADOX*

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. 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.

5.3.1.6 *Cyclone*

Cyclone has a theoretical peak performance of 600 TFlop/s. It consists of 17 compute nodes and 16 GPU nodes and is the new production HPC system of The Cyprus Institute since its deployment in 2020. All nodes are equipped with two Intel Xeon Gold 6248 CPUs. The GPU nodes have 4 NVIDIA Tesla V100-SXM2 32GB each. The memory on the compute nodes is 96GB and on the GPU nodes is 192GB. The machine has a flash storage capacity of 135TB and is also connected to the 3PB HPC storage. A HDR100 Infiniband network connects all the system components.

5.3.2. Cloud generic services

5.3.2.1 FINKI Cloud

Openstack cloud deployed at the Faculty of Computer Science and Engineering, UKIM. Cloud infrastructure based on Openstack and is hosted on 15 Huawei servers, each with 128GB RAM and 20 HT CPU cores, totaling in 300 vCPU cores and 37TB SSD and 32 TB SAS storage. The system is in production from 2017 as a National cloud system. The connectivity to 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 if fully integrated with the eduGAIN / NI4OS AAI, enablig seamles access to virtulizes computing, network and storage resources.

5.3.2.2 AVITOHOL cloud

The Avitohol Cloud service allows users to launch virtual machines on servers from the Avitohol supercomputer. It allows user groups to launch long running virtual machines with substantial flexibility. It is used by diverse research groups with needs for both advanced computing and data storage.

5.3.2.3 ICIPRO

ICIPRO offers Infrastructure as a Service services (IaaS) to Public Sector beneficiaries that need flexibility, modularity, dynamics and access to state of art technologies. Tenants can self-provision scalable Windows Server and Linux virtual machines from a gallery of predefined images.

5.3.2.4 GCloud.ge

GRENA is offering cloud service - GCloud. Besides with standard IaaS (Infrastructure as a Service) platform, we offer wide range of applications that can be installed automatically within a few minutes. Additionally, we have integrated various useful services for developers to make their work easy.

5.3.2.5 ASNET-AM Cloud

ASNET-Cloud provides Infrastructure as a Service services (IaaS) to academia and stakeholders based on OpenStack middleware. Additionally, ASNET-Cloud provides customized virtual environments to researchers and end-users to make their work easy. The users launch virtual machines (1-64 CPU cores) via dashboard distributed in the following three zones (596 CPU cores in total):

Zone 1: 160 CPU cores, 1TB RAM, 29TB storage

Zone 2: 320 CPU cores, 1TB RAM, 3.2TB storage

Zone 1: 116 CPU cores, 0.4TB RAM, 1.9TB storage

5.3.2.6 RASH Cloud

UCloud is cloud services offered by RASH for Albanian universities. Developed with Openstack and other open source technologies, it is simple to use and to configure.

5.3.2.7 RENAM Scientific Cloud

Provide computing resources running virtualized servers on the cloud for research, computational experiments, education, training, testing IT solutions and preparing services for further use in the activities of research and educational organizations.

5.3.2.8 UoM Cloud

Provide computing resources running virtualized servers on the cloud for research, computational experiments, education, training, testing IT solutions and preparing services for further use in the activities of research and educational organizations.

5.3.3.　　Storage and other generic services

5.3.3.1 ARIS Archival Service

Data archiving is the practice of moving data that is no longer being used or are being used on a less frequent fashion into a separate storage device. It is a single set, or a collection of historical records specifically selected for long term retention and future reference. The Archival service is based on the tape-based tertiary storage of the ARIS HPC system.

5.3.3.2 Data analysis service

Data analysis service 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 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.

5.3.3.3 Simple storage service

The Simple Storage Service (SSS) is a secure data storage service provided to researchers for storing and sharing research data as well as keeping it synchronized across different devices. SSS is functionally similar to Dropbox, Office 365, or Google Drive. Files are stored in conventional directory structures, accessible via WebDAV. User files are encrypted during transit and can be synchronized with local clients running Windows, macOS, or various Linux distributions. The service is based on the NextCloud platform.

5.3.3.4 Data discovery service

The service provides for flexible searching for data discovery. This is a powerful dataset management system that provides publishing, sharing, searching and can use almost any type of data and metadata. It provides users with ability to publish metadata about their data and to find relevant datasets based on metadata information. The service is open and multidisciplinary.

5.3.3.5 *RENAM Storage Service*

RENAM Storage Service allows research community members to keep and sync research data on Dell 740 servers with FreeNAS software. Available services: FTP, WebDAV, NFS. Can be used as network storage for VMs backup, exchange and archive information.

5.3.3.6 *EOSC RoP Legal & Ethics Compliance Tool (RoLECT)*

RoLECT aims to help in addressing the need of researchers to publish in FAIR/open modes. It targets open and FAIR assessors' needs in terms of IPR, ethics and data protection compliance both at the policy and legal level and aligned with the EOSC governance. RoLECT aims at providing an aggregated procedure for legal and ethics compliance by integrating a set of model procedures including: model procedures for copyright acquisition, management and dissemination policies; model copyright clearance processes, documentation and tools; model data protection (GDPR compliant) processes, consent forms and data sharing agreements; decision support trees for data protection policies; model IPR and data protection documentation. The intended use of the tool is to provide an aggregated guided assessment for EOSC Rules of Participation (RoP) focusing on legal and ethical aspects of compliance. Targeted users may be service providers, researchers and research organizations. The RoLECT platform will eventually evolve to automatically check the validity of the provided resources for at least the obligatory steps of the assessment. RoLECT is fully operational and stable and will be aligned with the forthcoming input from the EOSC RoP Task Force.

5.3.3.7 *License Clearance Tool (LCT)*

LCT aims to facilitate and automate the clearance of rights (copyright) for datasets, media, software and other content before they are released under an open license or stored at a publicly trusted FAIR repository. It helps in the certification of datasets and other outputs in terms of license compatibility analysis and selection as well as other related constraints. LCT helps in open research data management and certification for data repositories, aligning with activities and results of INFRAEOSC-5c. LCT check for equivalence, similarity and compatibility between licenses if used in combination, particularly for derivative works. In one scenario, the user (data manager) aims to find an appropriate open-source license for the set of the elements with separate licenses or to select a license for derivative work based on the content or components with various licenses. In the second scenario, the user declares the desired out-license and verifies the compatibility of the existing in-licenses with the derivative work. The LCT user may register or use it as a guest. The data provided by the guest is not preserved after the report is produced. For the authenticated user, the content and all parts of the procedure are kept in the licensing clearance history and user history, allowing to resume the work or share the work and data with others.

5.3.3.8 *RePol - Repository Policy Generator*

RePol – Repository Policy Generator is an open-source web application that guides users through the process of defining policies for repositories and web-based services. It helps in defining and maintaining comprehensive and clear repository and privacy policies. Generated privacy policies are suitable for any kind of service. RePol uses a step-by-step wizard and self-explanatory forms to guide users through the process. By choosing options in a form, users shape a policy document with predefined policy clauses formulated in line

with the current best practice. The resulting policy document may be downloaded as an XML file, additionally customized, and integrated into the service or repository. The collected data and the key elements of the generated policy are provided in a machine-readable format. This allows for an automated interpretation of created policies and extraction of repository-level metadata for inclusion in registries, catalogues and various operational and data discovery tools. The resulting policy document may be downloaded, additionally edited, and integrated into a repository. The main purpose of this extensible web application is to generate policies for repositories, but it can be configured to generate any other type of document, due to the versatile nature of its configurable forms and FreeMarker templates. RePol is a Java web application (using EE Web API 7.0) built upon Java Server Faces 2.2 framework, with PrimeFaces 7 components. It generates documents using FreeMarker 2.3 library.

5.3.3.9 *IoT Cloud Platform*

IoT Cloud Platform is a cloud application for allowing users to collect data from internet-connected hardware, and visualise it in near real-time. Collected data will be able to be analysed online in Octave or R programming language (discovering relationships, patterns and trends in data), whereby users will be able to use a pre-written algorithm or develop a new code directly in the web browser. IoT platform will support a variety of hardware devices such as Arduino and Raspberry Pi platforms, mobile devices, PCs, etc. Besides the web interface, an API will be provided, so that the data can be uploaded, stored and accessed by a third party.

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

5.4.1.   AGORA Resource Portfolio Management Tool

AGORA is a tool for managing a "service portfolio". It is addressed to the management board of an organization, to oversee all services, tools and products that it either uses internally or provides to its customers. Interrelationships of services and possible external dependencies are also described. The tool automatically generates a list of services available to customers. AGORA operates under an umbrella organization whose members are various academic institutions, as is the case with the European project EOSC-HUB. Its users are representatives of Institutions that record through this tool the services they have developed and provide. The tool supports the detailed recording of a service, including a) its available versions, b) the level of maturity, and c) user support. It also supports the recording of so-called service components. The architecture of the tool is based on the FitSM protocol and implements the EOSC profiles. Institutions are listed as service providers by the tool administrator. Ordinary users are then registered in the system and identified using the federal AAI and SAML. The administrator undertakes their assignment with the institutions – services. Then a user can register a service on behalf of his institution. From the logs in the tool emerges a list of available services, which is posted on the page of the umbrella operator. This list is aimed at researchers looking for available services.

5.4.2.   ARGO monitoring engine

The ARGO Monitoring Engine supports monitoring of status, availability, and reliability of services provided by infrastructure in NI4OS-Europe. It generates reports using customer-defined profiles such as for SLA management, operations, etc.

5.4.3.   Accounting system

The accounting service collects, analyzes and then provides information about the usage of services for example HPC usage, storage data, virtual machines data. The service provides information only related to NI4OS/EOSC services and is available to the project management team and the service administrators.

5.4.4.   Configuration management database

The configuration management database is a component of the monitoring system in NI4OS-Europe pre-production environment. It is based on the well-established GOCDB platform. The database maintains the central registry for e-infrastructure topology.

5.4.5.   NI4OS-Europe Login

The NI4OS-Europe Login service is the Authentication and Authorization Infrastructure (AAI) for the NI4OS-Europe infrastructure. It enables researchers to securely access and share common resources and services using identities from eduGAIN, the global network of academic identity federations. For users coming from the industry or citizen scientists who may not have access to eduGAIN, the NI4OS-Europe Login service supports additional trusted authentication providers, such as social networks, community identity providers and other platforms that can provide federated user identities. NI4OS-Europe Login allows connecting services using popular protocols, such as OpenID Connect and SAML, to securely authenticate and identify their users, organise them in groups, assign them roles and centrally manage access rights for accessing protected resources.

5.4.6.   NI4OS-Europe training and webinar portal

You will find training materials related to the training events organized by the NI4OS-Europe project, but also self-paces online courses on various topics related to the European Open Science Cloud.

5.4.7.   NI4OS-Europe helpdesk

The NI4OS-Europe helpdesk is the main tool of the project's support system. All stakeholders within the project use the helpdesk: resource users, resource owners, and project management. It supports ticket management, including ticket creation, assignment/reassignment to the appropriate unit, escalation, closing, searching, etc. The helpdesk classifies tickets per support unit and notifies relevant actors about ticket changes. The helpdesk is based on OTRS technology and can be easily integrated with external resources.

5.4.8.   NI4OS-Europe on-boarding dashboard

The main aim of the NI4OS-Europe as on-boarding dashboard is to provide a better and more user-friendly view of the resources/providers registered within the resource portfolio management system (AGORA catalogue) and track the progress of the providers and resources on-boarding process. The dashboard offers two main views: providers and

resources. From the providers' view, users can access the details of each of the registered providers as well as the on-boarding status of each provider. By selecting a particular provider, its detailed information page is shown. Once we collect all information marked as required within the AGORA, the provider is marked as on-boarded, while those who have not yet completed their profiles have the candidate/registered status. The resources view offers a list of all resources. It shows the basic information in the list, which can be extended by selecting a provider of interest. Resources in the list are categorized into four main categories: repositories, thematic services, generic services, and core services. A resource is tagged as on-boarded once all the required data is entered in the AGORA system. The dashboard helps the NI4OS-Europe on-boarding team spot the missing information and react in time. The dashboard relies on AGORA API, which has been optimized in the previous period to improve the dashboard's performance. This is done by decreasing the amount of information sent via API, so the API's endpoints return a fraction of the data available in the catalogue (only information that the dashboard needs to show).

6. Usage of the resources in the open call

NI4OS-Europe offers a broad set of generic, thematic services as well as repositories in the European region, with a special focus on the scientific communities of Life Sciences, Climatology, Digital Cultural Heritage, and Computational Physics. Generic services are in the areas of Compute resource provisioning (HPC and Cloud), Storage, and Data services provisioning. Repositories provide access to datasets, Software, as well as scientific manuscripts.

N4OS-Europe initiated an open call on the 4th of April 2022, which was addressed to scientists and researchers that work in academic and research institutions in EU members and associated countries. Via the open call, NI4OS-Europe provided possibilities for scientists to access the advanced services on-boarded to EOSC (thematic, generic, and repositories). The services provided have been tested in the use-cases run by scientific teams that were assigned to this task. The services were then fine-tuned according to the demonstrators' feedback and further tested via the open call. Access to the underlying computational resources (generic services) has been awarded for a maximum period of 6 months, while access to underlying storage resources can be provided for up to 8 months.

As a part of the open call preparation activities, in collaboration with resource providers we dedicated computational power to be offered in the call. Here, we primarily focused on generic services since a significant number of thematic services and repositories have a free access mode. As a result, we allocated 6.2 million CPU-core hours, 402,000 GPU-card hours, and 10,000 Xeon Phi-card hours at the ARIS, Avitohol, Leo HPC, Isabella, and PARADOX clusters. In addition, to support projects oriented to Cloud resources, we allocated in total 364 virtual machine cores at the FINKI, GCoud.ge, and ICIPRO cloud sites. Although the storage resources were by default provided at each HPC/Cloud site, additionally, we allocated 10 TB per project at the archival service provided by GRNET for archiving purposes and 100 GB per project at the simple storage service provided by IPB for results sharing.

Table 2: Thematic services requested during the NI4OS-Europe open call grouped per scientific community

Life Sciences	Climate	DCH	Computational Physics
ChemBioServer	LAS	CHERE	Gaussian API
DREAMM	OMApp	ClowderDCH	Schrodinger API
FEPrepare	RS2C		
Ingredio	SciRoHub		
NanoCrystal			
OpenBioMaps			
Reduce and Visualize Gene Ontology			
Eeghub.ge			

Throughout the Open Call, 21 project proposals have been received and after a technical and light scientific evaluation 20 of those have been accepted to be provided the required resources. Namely, from those 20 proposals, 5 belong to the thematic community of Life Sciences, 2 in Climate Research, 3 in Digital Cultural Heritage, 8 in Computational Physics and 2 in other fields.

From the 20 projects accepted to access NI4OS-Europe services, 16 projects required access to HPC resources, so we distributed 3.2 million CPU-core hours, 120 kGPU-card hours, and 50 TBs of storage space from the pool of allocated resources. In addition, 19 virtual machines with 166 virtual machine cores were created, and 13 TBs of storage space at the archival service was provided. Access to the following thematic services was granted to the open call projects: Gaussian API, Schrödinger API, OMApp, CHERE, Digital Library, and LCT. The most demanding repository in the open call was the NI4OS-Europe repository, which is the main general-purpose storage service that holds regional community datasets. All projects accepted during the open call are also granted full support while using provided resources. Resource providers directly offer such support as a part of WP5 activities via the project's helpdesk.

7. Conclusions

One of the key EOSC principles is to integrate the exiting pan-European resources, providing the research communities simplified access to data, resources and services, enabling them environment for scientific excellence. This key principle was embedded in the NI4OS-Europe project as one of the most important pillars, addressing the resources and their providers in the SEE region.

Due to a significant number of resources and providers within the NI4OS-Europe project, we developed the pre-production environment to support the on-boarding procedure using the approach based on the regional catalogue. The pre-production environment provides the tools for the services providers to integrate and validate their resources, before making them available for the EOSC marketplace. Within the environment, we deployed the project's catalogue based on the AGORA service portfolio management tool at the very beginning of the project. The catalogue relies on EOSC profiles specification that ensures compatibility with resources available in the EOSC marketplace. In addition, to be able to monitor the performance of the registered resources, we deployed a set of fully EOSC-compatible operational tools within the pre-production layer and performed integration of these tools with the available resources. Finally, we integrated our catalogue with the EOSC marketplace via a dedicated API to expose regional services to pan-European researchers. Practically, once a resource is marked as on-boarded within the NI4OS-Europe catalogue, it is automatically published as on-boarded into the EOSC marketplace.

We categorised resources within the catalogue into repositories, thematic services, generic services, and core services. Up to now, we have registered 39 repositories, 33 thematic, and 23 generic services within the AGORA catalogue. Out of these, 17 repositories, 18 thematic, and 18 generic services are fully on-boarded and published into the EOSC marketplace, while the others are in the process of on-boarding. These resources have been provided by 58 registered providers, out of which 53 are fully on-boarded. Comparing these numbers with the KPI defined in the project's proposal (20 generic service instances, 20 thematic services, and 15 repositories), we can conclude that we attract much more than planned resource providers to initiate EOSC integration progress, while the amount of fully on-boarded resources is very close to the goal and will be reached within the project's lifetime.

Finally, to assess the scientific value of the registered resources, we can use the data from the open call organised in April 2022, through which all registered services were offered to researchers. As a result, we granted to 20 accepted projects 3.2 million CPU-core hours, 120 kGPU-card hours, 166 virtual machine cores, 50 TBs of local storage space, and 13 TBs of storage space at the archival service. Since a large number of thematic services have a fully open access policy, we received a smaller number of requests for this type of service. All projects accepted during the open call are also granted full support while using provided resources.