D5.3 Periodical assessment of Data space services

|  |  |
| --- | --- |
| **Lead partner:** | EGI Foundation |
| **Version:** | 1 |
| **Status:** | Final |
| **Dissemination Level:** | Public |
| **Keywords:** | Virtual Access, Data Spaces, EGI-ACE |
| **Document Link:** | <https://documents.egi.eu/document/3809> |

|  |
| --- |
| **Deliverable Abstract** |
| The deliverable provides metrics and assessments of the 13 EGI-ACE Data Space services provided under the Virtual Access (VA) mechanism in WP5. |

**COPYRIGHT NOTICE**



This work by parties of the EGI-ACE consortium is licensed under a Creative Commons Attribution 4.0 International License. (<http://creativecommons.org/licenses/by/4.0/>).

EGI-ACE receives funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no. 101017567.

**DELIVERY SLIP**

|  |  |  |
| --- | --- | --- |
| Date | Name | Partner/Activity |
| From: | Giuseppe La Rocca | EGI Foundation / WP5 |
| Moderated by: | Sjomara Specht | EGI Foundation / WP1 |
| Reviewed by: | Enol Fernández Gergely Sipos | EGI Foundation EGI Foundation |
| Approved by: | SDS |  |

**DOCUMENT LOG**

|  |  |  |  |
| --- | --- | --- | --- |
| Issue | Date | Comment | Author |
| v.0.1 | 18/01/2022 | Table of content | M.Krakowian, H. Bui |
| v.0.2 | 28/03/2022 | Drafted initial content | G. La Rocca |
| v.0.3 | 29/04/2022 | Incorporated feedback from reviewers | G. La Rocca |
| v.1 | 04/05/2022 | Final | G. La Rocca |

**TERMINOLOGY**

<https://confluence.egi.eu/display/EGIG>

**Contents**

[Executive summary 5](#_Toc102798029)

[1. Introduction 7](#_Toc102798030)

[1.1 Installations 7](#_Toc102798031)

[1.2 Metrics definition 12](#_Toc102798032)

[**2** **Installations** 13](#_Toc102798033)

[2.1 WeNMR: A worldwide e-Infrastructure for NMR spectroscopy and Structural biology 13](#_Toc102798034)

[2.1.1 Metrics 14](#_Toc102798035)

[2.1.2 Assessment 21](#_Toc102798036)

[2.2 Virtual Imaging Platform (VIP) 22](#_Toc102798037)

[2.2.1 Metrics 23](#_Toc102798038)

[2.2.2 Assessment 24](#_Toc102798039)

[2.3 OpenRiskNet/NanoCommons Virtual Environment 25](#_Toc102798040)

[2.3.1 Metrics 27](#_Toc102798041)

[2.3.2 Assessment 27](#_Toc102798042)

[2.4 useGalaxy.eu 28](#_Toc102798043)

[2.4.1 Metrics 29](#_Toc102798044)

[2.4.2 Assessment 30](#_Toc102798045)

[2.5 OPENCoastS 30](#_Toc102798046)

[2.5.1 Metrics 32](#_Toc102798047)

[2.5.2 Assessment 32](#_Toc102798048)

[2.6 ENES Data Space 34](#_Toc102798049)

[2.6.1 Metrics 35](#_Toc102798050)

[2.6.2 Assessment 36](#_Toc102798051)

[2.7 PROMINENCE 37](#_Toc102798052)

[2.7.1 Metrics 39](#_Toc102798053)

[2.7.2 Assessment 39](#_Toc102798054)

[2.8 LOFAR Science Products 40](#_Toc102798055)

[2.8.1 Metrics 41](#_Toc102798056)

[2.8.2 Assessment 42](#_Toc102798057)

[2.9 SeaDataNet WebOcean Data Analysis 43](#_Toc102798058)

[2.9.1 Metrics 44](#_Toc102798059)

[2.9.2 Assessment 45](#_Toc102798060)

[2.10 EMSO ERIC data services 45](#_Toc102798061)

[2.10.1 Metrics 46](#_Toc102798062)

[2.10.2 Assessment 47](#_Toc102798063)

[2.11 GBIF Cloud data space 48](#_Toc102798064)

[2.11.1 Metrics 50](#_Toc102798065)

[2.11.2 Assessment 51](#_Toc102798066)

[2.12 Disaster mitigation and agriculture 52](#_Toc102798067)

[2.12.1 Metrics 53](#_Toc102798068)

[2.12.2 Assessment 53](#_Toc102798069)

[2.13 OPERAS Metrics service and Certification service 54](#_Toc102798070)

[2.13.1 Metrics 56](#_Toc102798071)

[2.13.2 Assessment 56](#_Toc102798072)

[3 Satisfaction 58](#_Toc102798073)

[3.1 The WeNMR Thematic Services 58](#_Toc102798074)

[3.2 The Virtual Imaging Platform 58](#_Toc102798075)

[3.3 The ENES Data Space 59](#_Toc102798076)

[3.4 EMSO ERIC Data Service 59](#_Toc102798077)

[4 Service Orders 60](#_Toc102798078)

[**Appendix I - Status of the WP5 integration activities** 61](#_Toc102798079)

Executive summary

This report provides an assessment at M15 of the 15 WP5 Thematic Service installations from 19 providers provided by the EGI-ACE project under the Virtual Access (VA) mechanism (see Figure 1). The Thematic Services of WP5 contribute to EGI-ACE Key Exploitation Result 3 - Research Data Spaces and Processing Tools.

The assessment was made based on the metrics collected by the 15 WP5 services during the three periods of observations covering the following three periods: M01-M05, M06-M10 and M11-M15. The EOSC Compute Platform, delivered by WP3, WP4 and WP6 is playing a fundamental role to deliver infrastructure and platform services for the scalable and open delivery of the 15 Data Space installations shown in Figure 1.

Diagram

Description automatically generated

Figure 1 - The WP5 Data Spaces and Thematic services landscape (after 15 months)

The Data Spaces and the Thematic services of WP5 in total served **65,340** users in 15 months, representing **65.98%** increase compared to the 15 months preceding the project start. The EOSC Compute Platform was instrumental to this scale up and delivered **10,553,094** Cloud CPU hours, more than **180** TB storage, **23,393** GPU hours, to the WP5 services. During the first part of the project we also registered more than **42M** of downloads and reached more than **50** countries worldwide.

By looking at the usage across the disciplines, the following can be observed:

* Within the **Health and Medicine** domain:WeNMR, VIP and UseGalaxy.eu served 57,444 users from structural biology, medical imaging and bioinformatics.
* Within the **Climate research** domain: OPENCoastS and ENES served 42 users[[1]](#footnote-1) from the Climate Change scientific domain. The assets offered by the EOSC Compute Platform allows scientists to run forecast simulations with the OPENCoastS service and predict a vast array of coastal dynamics variables. The ENES Data Space delivers a cloud-enabled data science environment for climate data analysis. The first release of the ENES Data Space was rolled out into production in Q3 2021.
* Within the **Energy and Physical Sciences** domain: PROMINENCE and LOFAR Science Products served 43 users from the Fusion and Astronomy domains. The EOSC Compute Platform contributed to open up the processing capabilities of the LOFAR Science Processing Data Space to a wider community of astronomers and develop a service that in the PROMINENCE Data Space will be of use to run significant modelling and use existing experimental data to perform validation.
* Within the **Environmental Sciences** domain: SeaDataNet WebOcean Data Analysis, EMSO-ERIC data services, GBIF Cloud data space and the Disaster Mitigation Agriculture data spaces served 4,642 users from the Environmental scientific domain. The EMO-ERIC Data Space reached pre-production level in 2021 and served 1,977 users/communities.
* Within the **Social Sciences and Humanities** domain: the OPERAS Metrics service and Certification service was on-boarded in EOSC Portal in Q1 2022.

To promote the uptake of new and existing services installations, 30 domain-specific training and dissemination activities were organized during the first part of the project. These activities played an important role to reach a wider user base in Europe as demonstrated by the metrics reported during the three periods of observations. Overall, these domain-specific training and dissemination events were attended by 1,703 participants.

In the second period:

* New releases of the ENES Data Space will be available every 5 months.
* A first release of the SeaDataNet WebOcean Data Analysis Data Space is expected during Q2 of 2022.
* The iCOMCOT science portal from the Disaster Mitigation and Agriculture Data Space is planned to be on-boarded in the EOSC Portal for Q4 of 2022.
* The EMSO ERIC Data Space is moving from the pre-production to the production level in 2022.
* The full integration of the production OPERAS Metrics service and Certification service in the EOSC Compute Platform is expected during Q3 of 2022.

Additional dissemination activities will be organized also during the second part of the project to promote the services uptake. More specifically, the following events have been already planned by the WP5 partners:

* The role of OPENCoastS+ on European Digital Ocean Twins initiative, April 2022.
* Tutorial on the Virtual Imaging Platform - Applications as a Service and Beyond, June 2022.
* OPENCoastS+: an EOSC-powered service for on-demand prediction of coastal water quality, July 2022.
* LOFAR Science Products and PROMINENCE Data Spaces will be presented in dedicated webinars in 2022.

# Introduction

Virtual Access (VA) is financial instruments to reimburse the access provisioning costs to access providers. This instrument is provided by the European Commission to increase the sharing of research infrastructures and services that otherwise would not be available to international user groups.

In VA, the services – also called “installations” – must be made available ‘free of charge at the point of use’ for European or International researchers. VA access is open and free access to services through communication networks to resources needed for research, without selecting the researchers to whom access is provided.

Virtual Access to services of the EGI-ACE catalogue applies to the following four categories:

1. Infrastructure Services WP3 - the Cloud Compute (IaaS) and High Throughput Compute services of the EGI portfolio supported by a set of 16 datacentres from the EGI Federation. The enabling components that support the Cloud Compute service: AppDB, for resource discovery and software catalogue; Dynamic DNS, for user-managed DNS provision of domain names for VMs and services running on the e-Infrastructure; and Infrastructure Manager (IM) for the basic orchestration of IaaS resources.
2. Platform Services WP4 - mature software tools offering generic capabilities to facilitate the usage of the underlying infrastructure for EOSC users and Data Spaces.
3. Federated data spaces WP5 - services provided by major European research collaborations, research infrastructures and research institutes, and are composed of mature software tools, datasets and services that offer science discipline specific processing and data analysis capabilities for EOSC users.
4. Federated Access Services WP6 – services providing secure access to other services and enabling large-data analysis workloads in the distributed infrastructure. Included services are delivered by major European research institutions using mature open-source software with already established user communities from multiple scientific disciplines.

This document provides Virtual Access metrics and assessment for WP5 - Federated data spaces.

## Installations

Within EGI-ACE project 15 installations are part of Virtual Access Work Package 5.

The status of the integration activities of the WP5 Thematic Services and Data Space installations, and the EOSC Compute Platform services (WP3, 4, 6 and 7) is shown in Appendix I.

The distribution of the EGI-ACE Data Spaces and Thematic Services by scientific disciplines and the number of users served per disciplines are reported in the Figures below:

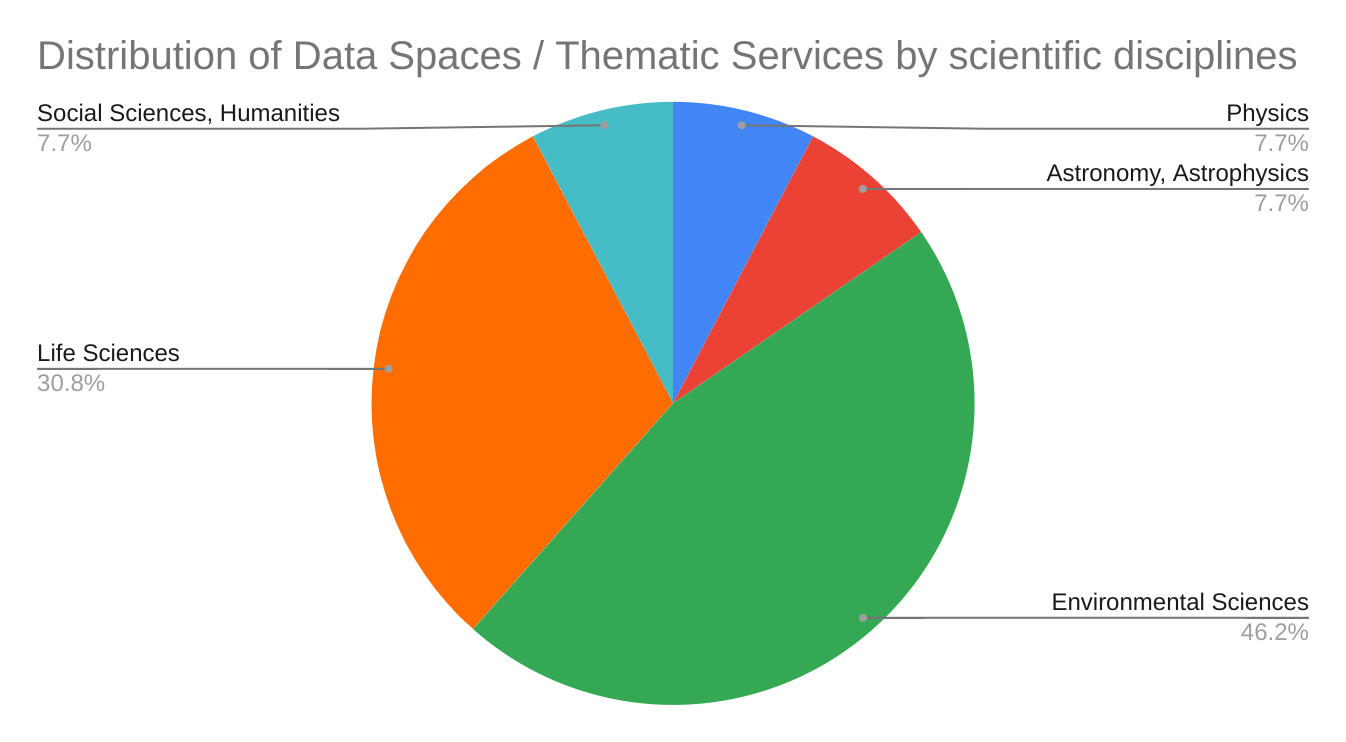


Figure 2 - Distribution of service installations across domains

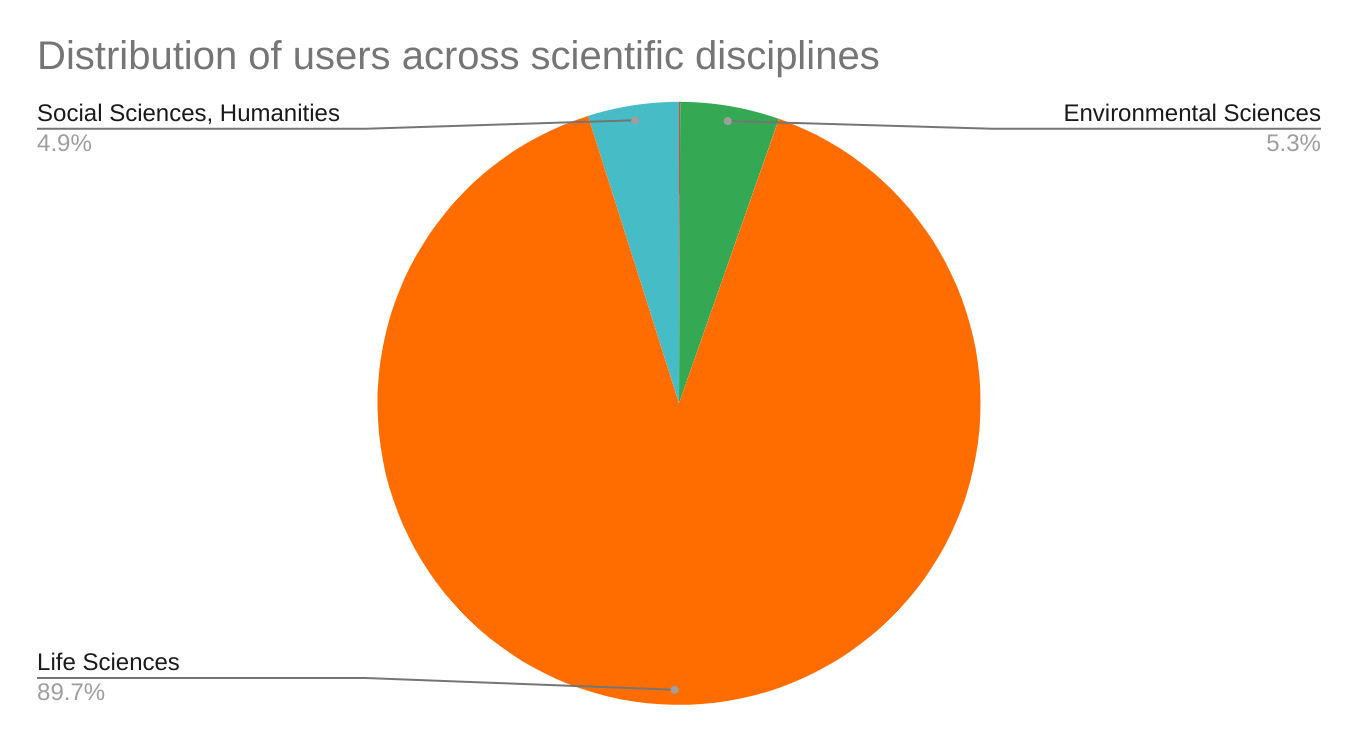
****

Figure 3 - Distribution of users per scientific disciplines

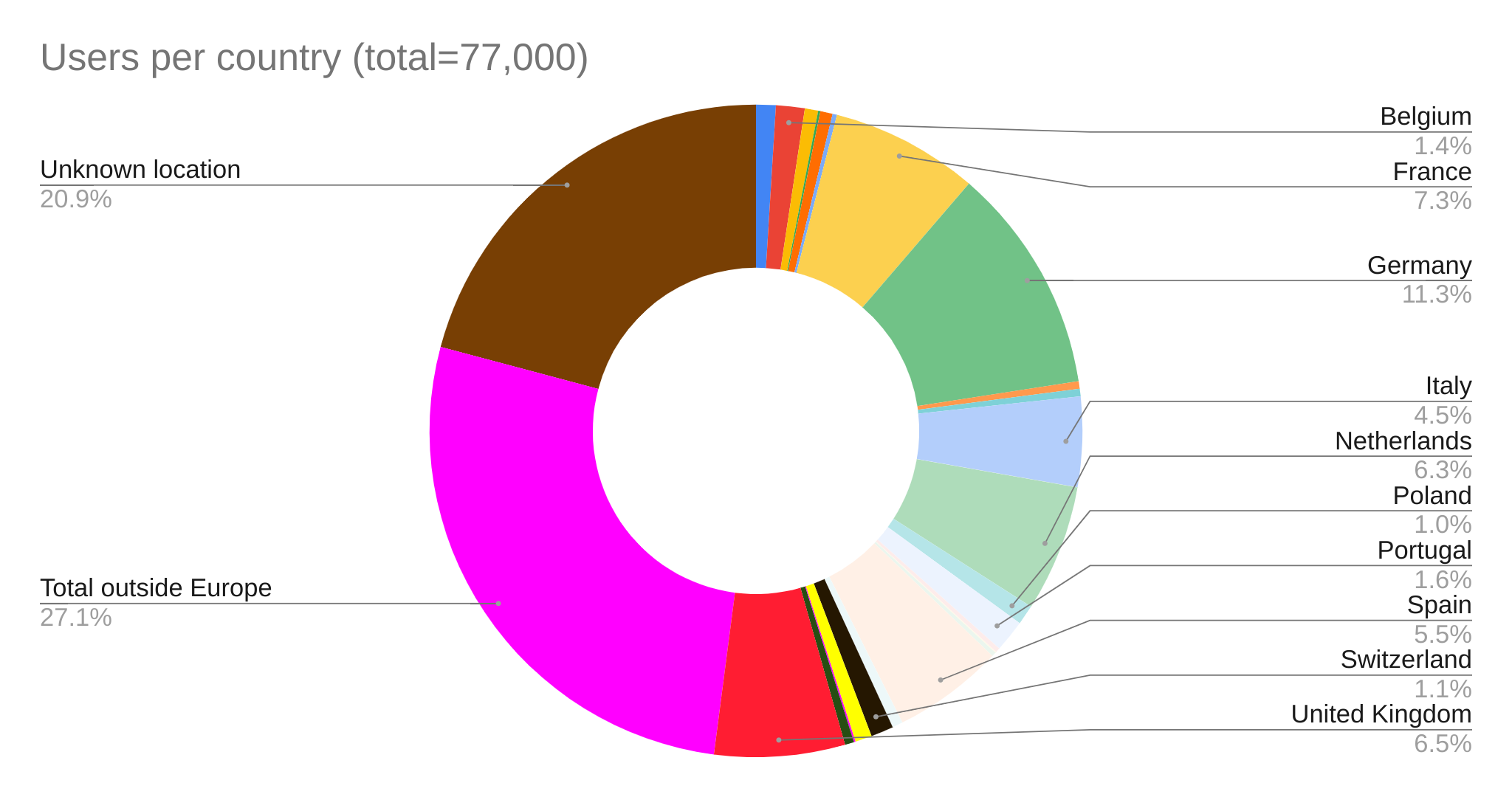
****

Figure 4 - Users distribution per country

The variation (in %) of the number of users per Data Space and/or Thematic service is summarised in the tables below.

Table 1 - Status of the Health and Medicine Data Spaces (T5.1)

|  | **WeNMR** | **VIP** | **useGalaxy.eu** | **OpenRiskNet/ NanoCommons Virtual Environment** |
| --- | --- | --- | --- | --- |
| **Metric** | *No. of new registered users* | *No. of registered users* | *No. of registered users* | *No. of registered users* |
| **15 months before the project** | 5,625 | 1,250 | 25,000 | 125 |
| **During M01-15** | 8,056 | 1,388 | 48,000 | 0 |
| **Variation (%)** | **+41.45%** | **+11%** | **+92%** | **N/A** |
| **Explanation** | Continuing the operation of the thematic portals.  Organized 9 events to promote the services. | Integrated 1 new application. Started to promote the platform. Organized 4 events to promote the service. | Integration with the EOSC Compute Platform is in progress. | Old service was revoked from the EOSC Portal, new service isn’t ready yet. |

Table 2 - Status of the Climate Data Spaces (T5.2)

|  | **OPENCoastS** | **ENES Data Space** |
| --- | --- | --- |
| **Metric** | *No. of registered users* | *No. of active users* |
| **15 months before the project** | 312 | 92 |
| **During M01-15** | 16 | 24[[2]](#footnote-2) |
| **Variation (%)** | **-94.87%** | **-97.82%** |
| **Explanation** | Limited opportunities for live dissemination and training events due to COVID restrictions.  Development activities during the first part of the project.  Organized 4 events to promote the service. | Focus on a significantly new release that became available towards the end of 2021. Promotion of the new platform started recently.  Organized 2 events to promote the service. |

Table 3 - Status of the Energy and Physical Sciences (T5.3)

|  | **PROMINENCE** | **LOFAR Science Product** |
| --- | --- | --- |
| **Metric** | *No. of users requesting access* | *No. of users requesting access to LOFAR* |
| **15 months before the project** | 3.75 | 0 |
| **During M01-15** | 2 | 41 |
| **Variation (%)** | **+46.67%** | **N/A** |
| **Explanation** | Increased pledged resources assigned to the Data Space.  Started the HPC pilot activities to run the JOREK non-linear MHD code. | First release rolled out into production in 2021 Q3.  Organized 3 events to promote the service. |

Table 4 - Status of the Environmental Sciences (T5.4)

|  | **SeaDataNet WebOcean Data Analysis** | **EMSO ERIC data services** | **GBIF Cloud Data Space** | **Disaster Mitigation and Agriculture** |
| --- | --- | --- | --- | --- |
| **Metric** | *No. of new registered users* | *No. of new users* | *No. of unique users* | *No. of new registered users* |
| **15 months before the project** | 125 | 250 | 10,038 | 187 |
| **During M01-15** | 0 | 1,977 | 6,858 | 0 |
| **Variation (%)** | **N/A** | **+690%** | **-46.36%** | **N/A** |
| **Explanation** | Integration of the Data Space with the EOSC Compute Platform is still in progress. | Integrated 1 new application. Started to promote the platform. | Integration in progress | Integration of the Data Space with the EOSC Compute Platform in progress. |

Table 5 - Status of the Social Sciences and Humanities (T5.5)

|  | **Operas Metrics service and Certification service** |
| --- | --- |
| **Metric** | *No. of registered publishers* |
| **15 months before the project** | 12 |
| **During M01-15** | 0 |
| **Variation (%)** | **N/A** |
| **Explanation** | Integration of the Data Space with the EOSC Compute Platform in progress. |

Following installations have been subject to change since the beginning of the project:

* 1 service has been revoked from WP5 from the EOSC Portal (OpenRiskNet/NanoCommons Virtual Environment). A new service will be on-boarded during the second part of the project.
* The integration plans of the two unfunded Data Spaces, including OpenRiskNet, and the Disaster Mitigation and Agriculture underwent a delay due to the lack of human effort being available during the first part of the project. The first service from the Disaster Mitigation and Agriculture Data Space, iCOMCOT, and the OpenRiskNet service are expected for 2022.

## Metrics definition

For each installation several metrics have been defined between the provider and WP5 leader, taking into account following categories:

* **Number of users** – depending on the nature of installation, number could be defined based on accounts (if registration was required) or number of unique IPs (if registration is not needed to benefit of the service).
* **Usage** – the goal of this metric is to report how much the service is used. This metric depended on functionality provided by the service.
* **Number and names of the countries reached** – the goal of this metric was to report how broadly the service is used and how the geographical coverage is changing with time.
* **Marketplace views** – the goal of this metric is to provide information about how often the service is being viewed by the potential customers.
  + This metric is not applicable to federation services due to the nature of the service. Federation services are enabling federation and are supporting delivery of customer facing services. Thus, cannot be ordered.
* **Marketplace orders** – the goal of this metric is to provide information about how often the service is being ordered via EOSC Marketplace.
  + This metric is not applicable to federation services due to the nature of the service. Federation services are enabling federation and are supporting delivery of customer facing services. Thus, cannot be ordered.

1. **Installations**
   1. WeNMR: A worldwide e-Infrastructure for NMR spectroscopy and Structural biology

|  |  |
| --- | --- |
| **Description** | The WeNMR services consist of a suite of web portals, providing user-friendly access to complex computational workflows and tasks. The WeNMR data analytics platform consists of a collection of user-friendly portals serving a community of over 16’000 users worldwide. The WeNMR services allow inexperienced and experienced structural biologists to use state-of-the-art software for their data analysis while benefiting from the EOSC computational infrastructure.  The services make use of high-throughput computing (HTC) resources, but some are also using GPGPU grid resources and cloud computing. The portals are already integrated with the EOSC AAI, present in the EOSC Portal and Marketplace, and use the EGI Check-in and DIRAC4EGI services to send ~10 million jobs per year to HTC resources.  WeNMR has been successfully serving the structural biology community for over 10 years now. The community shows a sustained growth with > 3500 new users per year. The WeNMR services are fully operational, all available under the EOSC portal and marketplace. |
| **Task** | 5.1 |
| **URL** |  |
| **Service Category** | Data Spaces and Analytics |
| **Service Catalogue** | <http://www.wenmr.eu/> |
| **Location** | Utrecht (NL), Florence (IT) |
| **Duration** | M01-M30 |
| **Modality of access** | Web interfaces |
| **Support offered** | The planned activities would encompass user support, training, and continuous operation of the various grid- and cloud-enabled web portals. |
| **Operational since** | Some of the services have been operational since June 2008 |
| **User definition** | A user is a person making use of at least one of our thematic services. All portals except one (FANTEN) do require user registration. For FANTEN, users are identified by their IP address for collecting the various metrics. |

### Metrics

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Metric name** | **Baseline** | **Define how measurement is done** | **Period 1 M01-M5** | **Period 2 M06-M10** | **Period 3 M11-M15** |
| *No of user runs submitted* | 27,500 | Internal logs of the service / accounting | 41,884 | 42,783 | 41,250 |
| *No of grid/cloud jobs submitted* | 1,440,000 | Internal logs of the service / accounting | 1,193,844 | 1,360,972 | 1,075,979 |
| HS06 CPU Time/Wall Time hours consumed by job submitted to grid/cloud resources | 17,500,000 grid / 75,000 cloud | Internal logs of the service / accounting | 18,168,871 (grid) / 103,233 (cloud) | 24,907,699 (grid) / 52,613 (cloud) | 25,938,182 (HS06 CPU Time hours (grid+cloud)) |
| No of countries reach | 110 | Internal logs of the service / accounting | 128 | 135 | 138 |
| Names of countries reach | Worldwide | Internal logs of the service / accounting | Afghanistan  Albania  Algeria  Argentina  Armenia  Australia  Austria  Azerbaijan  Bahrain  Bangladesh  Belarus  Belgium  Belize  Bolivia  Bosnia and Herzegovina  Brazil  Brunei  Bulgaria  Cameroon  Canada  Chad  Chile  China  Colombia  Comoros  Costa Rica  Croatia  Cuba  Cyprus  Czech Republic  Czechia  Côte d'Ivoire  Denmark  Dominican Republic  Ecuador  Egypt  El Salvador  Estonia  Ethiopia  Finland  France  Georgia  Germany  Ghana  Greece  Greenland  Guatemala  Hong Kong  Hungary  Iceland  India  Indonesia  Iran  Iraq  Ireland  Israel  Italy  Japan  Jordan  Kazakhstan  Kenya  Kuwait  Latvia  Lebanon  Lithuania  Luxembourg  Macao  Macedonia  Malawi  Malaysia  Maldives  Malta  Mexico  Morocco  Nepal  Netherlands  New Zealand  Nigeria  Niue  Norway  Oman  Pakistan  Palestine  Panama  Papua New Guinea  Paraguay  Peru  Philippines  Poland  Portugal  Puerto Rico  Qatar  Romania  Russia  Rwanda  Réunion  Saint Lucia  Saudi Arabia  Senegal  Serbia  Singapore  Slovakia  Slovenia  South Africa  South Korea  Spain  Sri Lanka  Sudan  Sweden  Switzerland  Taiwan  Thailand  Timor-Leste  Togo  Tunisia  Turkey  Tuvalu  Uganda  Ukraine  United Arab Emirates  United Kingdom  United States  Uruguay  Uzbekistan  Venezuela  Virgin Islands  Vietnam  Zimbabwe | Afghanistan  Albania  Algeria  American Samoa  Argentina  Armenia  Australia  Austria  Azerbaijan  Bahrain  Bangladesh  Belarus  Belgium  Belize  Bolivia  Bosnia and Herzegovina  Brazil  Brunei  Bulgaria  Cameroon  Canada  Chad  Chile  China  Colombia  Comoros  Costa Rica  Croatia  Cuba  Cyprus  Czech Republic  Côte d'Ivoire  Denmark  Dominican Republic  Ecuador  Egypt  El Salvador  Estonia  Ethiopia  Finland  France  Gambia  Georgia  Germany  Ghana  Greece  Greenland  Guatemala  Haiti  Hong Kong  Hungary  Iceland  India  Indonesia  Iran  Iraq  Ireland  Israel  Italy  Japan  Jordan  Kazakhstan  Kenya  Kuwait  Laos  Latvia  Lebanon  Lithuania  Luxembourg  Macao  Macedonia  Malawi  Malaysia  Maldives  Mali  Malta  Mexico  Mongolia  Morocco  Myanmar  Namibia  Nepal  Netherlands  New Zealand  Nigeria  Niue  Norway  Oman  Pakistan  Palestine  Panama  Papua New Guinea  Paraguay  Peru  Philippines  Poland  Portugal  Puerto Rico  Qatar  Romania  Russia  Rwanda  Réunion  Saint Lucia  Saudi Arabia  Senegal  Serbia  Singapore  Slovakia  Slovenia  South Africa  South Korea  Spain  Sri Lanka  Sudan  Sweden  Switzerland  Taiwan  Thailand  Timor-Leste  Togo  Tunisia  Turkey  Tuvalu  Uganda  Ukraine  United Arab Emirates  United Kingdom  United States  Uruguay  Uzbekistan  Venezuela  Virgin Islands  Vietnam  Zimbabwe | Åland Islands  Afghanistan  Albania  Algeria  American Samoa  Argentina  Armenia  Australia  Austria  Azerbaijan  Bahrain  Bangladesh  Belarus  Belgium  Belize  Bolivia  Bosnia and Herzegovina  Brazil  Brunei  Bulgaria  Cameroon  Canada  Chad  Chile  China  Colombia  Comoros  Costa Rica  Croatia  Cuba  Cyprus  Czech Republic  Cote d'Ivoire  Denmark  Dominican Republic  Ecuador  Egypt  El Salvador  Estonia  Ethiopia  Finland  France  Gambia  Georgia  Germany  Ghana  Greece  Greenland  Guatemala  Haiti  Hong Kong  Hungary  Iceland  India  Indonesia  Iran  Iraq  Ireland  Israel  Italy  Japan  Jordan  Kazakhstan  Kenya  Kuwait  Laos  Latvia  Lebanon  Lithuania  Luxembourg  Macao  Macedonia  Malawi  Malaysia  Maldives  Mali  Malta  Mexico  Moldova  Mongolia  Morocco  Myanmar  Namibia  Nepal  Netherlands  New Zealand  Nigeria  Niue  Norway  Oman  Pakistan  Palestine  Panama  Papua New Guinea  Paraguay  Peru  Philippines  Poland  Portugal  Puerto Rico  Qatar  Romania  Russia  Rwanda  Reunion  Saint Lucia  Saudi Arabia  Senegal  Serbia  Singapore  Slovakia  Slovenia  South Africa  South Korea  Spain  Sri Lanka  Sudan  Sweden  Switzerland  Taiwan  Thailand  Timor-Leste  Togo  Tunisia  Turkey  Tuvalu  Uganda  Ukraine  United Arab Emirates  United Kingdom  United States  Ukraine  Uruguay  Uzbekistan  Venezuela  Virgin Islands  Vietnam  Zimbabwe |
| *No of new registered users* | 4,500 | Internal logs of the service / accounting | 2,660 | 2,720 | 2,676 |

### Assessment

The WeNMR Thematic Services, supported by the UU[[3]](#footnote-3) and CIRMMP, have been fully operational for many years by now with demonstrated impact and usage. All the thematic services are available in the EOSC Portal and Marketplace[[4]](#footnote-4) (under the Software category). Access to the computing and storage resources to support the operation of the Thematic Services was formalized with EGI through a Service Level Agreement (SLA)[[5]](#footnote-5) which has been extended until June 2023. A total of 50+ Million (HTC) CPU/h (opportunistic access), 500+ cloud CPU cores and 60TB of storage are provisioned to allow the structural biologist community to have a transparent access to computing facilities. This capacity allocation is officially supported by SARA-Matrix and NIKHEF (NL), TW-NCHC (TW), NCG-INGRID-LP (PT), INFN-PADOVA-STACK and INFN-LNL-2 (IT), CESNET-MCC (CZ), IFCA-LCG2 and CESGA (ES) and UA-BITP (UA).

The integration with the EOSC Compute Platform, and the use of resources, are instrumental for continuing to offer free access to the WeNMR Thematic Services via web portals. The benefits that EGI-ACE is bringing to the WeNMR Thematic services are clearer whether we consider the % of increment of the VA metrics reported at the beginning of the project and after 12 months. More specifically:

* No. of user runs submitted (before the project start): 34,375[[6]](#footnote-6)
* No. of user runs submitted at M15: 41,250, with an increment of +20%
* No. of new registered users (before the project start): 5,6953
* No. of new registered users at M15: 2,660+2,720+2,676=8,056, with an increment of +41.45%

Several outreach and training events were organized by the University of Utrecht (UU) and the Consorzio Interuniversitario Risonanze Magnetiche di Metallo Proteine (CIRMMP)[[7]](#footnote-7) to promote the uptake of the services suite across the structural biologists community. A total of 12 training events were organized during the first part of the project. As shown by the metrics collected during the three periods of observation (see table above), outreach and training activities actively contributed to promote the uptake of the WeNMR Thematic Services to reach a wider user base in Europe and worldwide. Most of the metrics have already met the target baseline. We expect to further improve this positive trend also during the second part of the project.

As part of the dissemination activities, a dedicated case study[[8]](#footnote-8) was also published in the EGI web site. From a technical point of view, during the second part of the project, special assistance will be devoted to move from X.509 certificates to token-based authentication and support the integration with the EGI Workload Manager[[9]](#footnote-9) service. During the second part of the project the Data Space will also investigate the possibility of integrating the HPC resources in the WeNMR portal.

## Virtual Imaging Platform (VIP)

|  |  |
| --- | --- |
| **Description** | VIP (Virtual Imaging Platform) is a web portal for the simulation and processing of massive data in medical imaging. VIP users can access applications as a service and significant amounts of computing resources and storage (provided by the biomed EGI Virtual Organisation) with no required technical skills beyond the use of a web browser. VIP is thus both a :  -- Service provider, in the sense that it provides users with applications as a service and various other service for FAIR data analysis through Boutiques (containers, publication to Zenodo, DOIs).  -- Consumer of resources, in the sense that applications available in VIP exploit HTC computing, storage and GPU resources provided by the biomed EGI VO.  Medical imaging applications have always been compute intensive. In the last few years, in addition to the usual CPU computing needs, GPU usage has become mandatory for the processing of (3D) medical data, as well as for efficient machine learning approaches such as deep learning. The service enables the life sciences medical imaging community to have a transparent access to such computing facilities, especially for collaborators with no specific computer science background. A typical use-case consists in :  -- Training phase: researchers build and train deep learning algorithms and models (GPUs needed)  -- Testing phase: medical doctors/specialists use these models to test their data (GPU or CPU) |
| **Task** | 5.1 |
| **URL** |  |
| **Service Category** | Data Spaces and Analytics |
| **Service Catalogue** | <https://www.creatis.insa-lyon.fr/vip/> |
| **Location** | INSA, Lyon, France |
| **Duration** | M01-M30 |
| **Modality of access** | Web interfaces and APIs |
| **Support offered** | Helpdesk, technical support will be provided for integration use cases. Training/workshops will be provided |
| **Operational since** | 2011 |
| **User definition** | A user from life sciences medical imaging community, more specifically to the medical imaging research community |

### Metrics

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Metric name** | **Baseline** | **Define how measurement is done** | **Period 1 M01-M5** | **Period 2 M06-M10** | **Period 3 M11-M15** |
| No of registered users | 1,000 | Internal logs of the service / accounting | 1,300 | 1,366 | 1,388 |
| No of use cases | 2 | Internal logs of the service / accounting | 1 | 0 | 1 |
| No of countries reach | 75 | Internal logs of the service / accounting | 81 | 81 | 82 |
| Names of countries reach | Worldwide | Internal logs of the service / accounting | Worldwide | Worldwide | Worldwide |

### Assessment

The main impacts of the EGI-ACE project on the Virtual Image Platform (VIP)[[10]](#footnote-10) Data Space for the first 15 months concern the integration of the EGI Check-in[[11]](#footnote-11) service, the user support with new applications and computing resources (GPUs), as well as a good dissemination campaign that should allow for the discovery and adoption of the VIP open services by new users. From a technical point of view, the integration of the Virtual Image Platform[[12]](#footnote-12) with EGI AAI Check-in has been completed and is now available in production. This integration has also led to a presentation[[13]](#footnote-13) at the ISGC conference in March 2022.

From user point of view, a new scientific application (the BRATS pre-processing pipeline[[14]](#footnote-14)) has been integrated in VIP for the Neuroimaging community. VIP also participated in the MSSEG-2[[15]](#footnote-15) scientific challenge and provided, through EGI ACE, the computing resources necessary (special GPU request) for the execution of the 31 pipelines integrated in VIP for the challenge. This event led to a presentation[[16]](#footnote-16) at the EGI conference in 2021.

Access to the computing and storage resources to support the operation of the Platform has been formalized with EGI through a Service Level Agreement (SLA)[[17]](#footnote-17) which has been extended until June 2023. A total of 440+ Million (HTC) CPU/h (opportunistic access), 350+ cloud CPU cores and 25TB of storage have been provisioned by 14 resources providers to allow the Life Sciences medical imaging community to have a transparent access to computing facilities. This capacity allocation is officially supported by BEIJING-LCG2 (CN), CESNET-MCC (CZ), CLOUDIFIN (RO), CREATIS-INRA-LYON, OBSPM, GRIF, IN2P3-CPPM and IN2P3-IRES (FR), INFN-BARI, INFN-CATANIA, INFN-FERRARA, INFN-PISA, and INFN-ROMA3 (IT), NCG-INGRID-PT (PT).

Outreach and dissemination activities have contributed to promote the Data Space and increase the metrics collected during the three periods of observation. A dedicated case study[[18]](#footnote-18) was published in the EGI web site and a webinar[[19]](#footnote-19) was organised in March 2022. For the second part of the project we expect to integrate additional scientific applications in the Platform and continue with the outreach and dissemination activities. The Data Space will also investigate the possibility of integrating the EGI Notebooks service in the platform.

During the last period of observation, the following % of increment was registered for the Data Space:

* No. of registered users (before the start of the project): 1,250[[20]](#footnote-20)
* No. of registered users at M15: 1,388, with an increment of +11%

## OpenRiskNet/NanoCommons Virtual Environment

|  |  |
| --- | --- |
| **Description** | OpenRiskNet operates a reference infrastructure consisting of 45 services grouped into seven categories: 1) Toxicology, Chemical Properties and Bioassay Databases, 2) Omics Databases, 3) Knowledge Bases and Data Mining, 4) Ontology Services, 5) Processing and Analysis, 6) Predictive Toxicology and 7) Workflows, Visualisation and Reporting. This infrastructure will be ported to the EGI-ACE cloud platform and offered to EOSC users by the project who will be able to test their functionalities and their applicability to their own specific study requirements, then to apply for additional EOSC resources, including but not limited to the EGI-ACE platform to setup and operate private environments to perform the actual risk assessments or safe-by-design studies. The risk assessment infrastructure will be further optimized to better integrate with other services of EGI-ACE, including AAI, HPC, Jupyter for making the provided data sources more visible and interlinkable with data from other relevant communities.  OpenRiskNet and NanoCommons provide concepts and guidelines for data management and sharing, specialized databases and software as well as a standardized cloud setup for the core infrastructure, and guidelines for the deployment of data and compute services on top of this core. With the latter, it is possible to set up virtual environments, in which the user can deploy the needed tools in a harmonized and interoperable way and execute workflows using Jupyter notebooks or visual workflow managers like Squonk developed by one OpenRiskNet partner. The service integration also included the development of workflows to support the case study work by automating complex tasks only achievable by the combination of multiple services. Additional services are being integrated by NanoCommons and external partners to complete the portfolio to allow full risk assessment of chemical compounds and nanomaterials, including tools for image analysis to predict nanomaterials properties or ecotoxicity, a range of QSAR models, tools for prediction of molecular initiating events and adverse outcome pathways and more. |
| **Task** | 5.1 |
| **URL** |  |
| **Service Category** | Data Spaces and Analytics |
| **Service Catalogue** |  |
| **Location** | Johannes Gutenberg Universität Mainz Germany (this will be replaced by an installation at EGI since the support is running out in 2021) |
| **Duration** | M01-M30 |
| **Modality of access** | Web interfaces |
| **Support offered** | Technical support for use cases, integration policy. Training, user documentation will be offered. |
| **Operational since** | June 2020 |
| **User definition** | Users for OpenRiskNet/nanoCommons services are typical industry and academic researchers, risk assessors, data managers, software developers |

### Metrics

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Metric name** | **Baseline** | **Define how measurement is done** | **Period 1 M01-M5** | **Period 2 M06-M10** | **Period 3 M11-M15** |
| No of new registered users | 100 | SERVICE STILL IN INTEGRATION PHASE | N/A | N/A | N/A |
| Workload of worker nodes | 0 | SERVICE STILL IN INTEGRATION PHASE | N/A | N/A | N/A |
| No of countries reach | 15 | SERVICE STILL IN INTEGRATION PHASE | N/A | N/A | N/A |
| Names of countries reach | European | SERVICE STILL IN INTEGRATION PHASE | N/A | N/A | N/A |
| No of new registered users | 100 | SERVICE STILL IN INTEGRATION PHASE | N/A | N/A | N/A |
| Workload of worker nodes | 0 | SERVICE STILL IN INTEGRATION PHASE | N/A | N/A | N/A |
| No of countries reach | 15 | SERVICE STILL IN INTEGRATION PHASE | N/A | N/A | N/A |
| Names of countries reach | European | SERVICE STILL IN INTEGRATION PHASE | N/A | N/A | N/A |

### Assessment

The OpenRiskNet/NanoCommons Virtual Environment Data Space is unfunded in EGI-ACE. This Data Space is supported by the University of Birmingham and Edelweiss Connect. The ambition of this Data Space was to increase the access and use of data, services and tools based on new approach methods supporting next generation risk and safety assessment, improve data quality and integrity, provide a reliable deployed infrastructure and services, and expose values of services to broader scientific communities.

The original plan was to move the OpenRiskNet risk assessment servicesin the EOSC Compute Platform. For supporting the porting of the OpenRiskNet risk assessment service, an initial capacity allocation composed of 32 vCPU cores, 64GB of RAM and 1TB of storage was enabled in the EOSC Compute Platform. The capacity allocation is supported by IFCA-LCG2 (ES) with dedicated Operational Level Agreement (OLA)[[21]](#footnote-21). Unfortunately, due to the lack of resources and budget allocated, the deployment activity was put on hold during the first part of the project.

Members of the Data Space have commenced deployment work in late March 2022, starting with a NanoCommons Transnational Activity aiming to deploy and evaluate several resources by June 2022 and use them in a virtual training. The plan for the next part of the project is also to deploy OpenTox and OpenRiskNet resources over the period April - August 2022, investigate the adoption of the Authentication and Authorization solution offered by EGI-ACE, and support a virtual summer school activity at the end of July 2022.

## useGalaxy.eu

|  |  |
| --- | --- |
| **Description** | The European Galaxy server (https://usegalaxy.eu) is the biggest Galaxy instance in Europe, and one of the biggest worldwide. This service provides access to underlying HPC and Cloud resources to more than 20.000 researchers. The service will make use of cloud compute, workload management and AAI services from EGI-ACE. The European Galaxy server is part of EOSC-Life, used by EOSC-Nordic and listed in the EOSC marketplace (https://marketplace.eosc-portal.eu/services/european-galaxy-server). In particular, UseGalaxy.eu provides: - compute and storage resource without any charge, - more than 2500 well-documented and constantly maintained tools, - 40.000 automatically built and tested containers, - 7 TB of reference data shared via CMVFS, - 250 GB quota per user (500 GB for ELIXIR members), - free registration, - Training Infrastructure as a Service (TIaaS). |
| **Task** | 5.1 |
| **URL** |  |
| **Service Category** | Data Spaces and Analytics |
| **Service Catalogue** | <https://usegalaxy.eu/> |
| **Location** | Freiburg (Germany) |
| **Duration** | M01-M30 |
| **Modality of access** | Web interfaces and APIs |
| **Support offered** | Service is open access upon registration. User documentation is in place. Technical support is provided upon requests |
| **Operational since** | 2013 |
| **User definition** | Individual researchers and communities in genomics, proteomics, metabolomics, ecology, climate-science, material-science, machine learning and many more |

### Metrics

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Metric name** | **Baseline** | **Define how measurement is done** | **Period 1 M1-M5** | **Period 2 M6-M10** | **Period 3 M11-M15** |
| No of user jobs | 200,000 | Public monitoring under https://stats.galaxyproject.eu | 20,000,000 | 34,318,670 | 44,000,000 |
| No of countries reach | 60+ | Public monitoring under <https://stats.galaxyproject.eu> | 80 | 90 | 92 |
| Names of countries reach | Worldwide | Public monitoring under <https://stats.galaxyproject.eu> | worldwide | worldwide | worldwide |
| No of registered users | 20,000 | Public monitoring under <https://stats.galaxyproject.eu> | 32,000 | 40,000 | 48,000 |

### Assessment

During the reporting period the useGalaxy.eu Data Space[[22]](#footnote-22) has started the integration with EGI Check-in[[23]](#footnote-23) to allow users to be authenticated using their Identity Providers. In addition, the EGI Datahub[[24]](#footnote-24) service was assessed to allocate computing capacities near data, and the EGI Workload Manager[[25]](#footnote-25) service was taken into consideration to implement a meta-scheduler in order to dispatch jobs both in the Pulsar network and in the resources procured by the EOSC Compute Platform. In terms of metrics, all the target baseline have been already met.

During the last period of observation, the following % of increment were registered for the Data Space:

* No. of registered users (before the start of the project): 25,000[[26]](#footnote-26)
* No. of registered users at M15: 48,000 new users, with an increment of +92%

## OPENCoastS

|  |  |
| --- | --- |
| **Description** | The OPENCoastS\_PLUS EOSC service will provide on-demand water quality hindcast and forecast simulations for the coastal region selected by each user. It builds on OPENCoastS, an on-demand circulation forecast service already integrated in EOSC through the EOSC-hub project. The improved version will include:  -- The capacity to forecast water quality variables, namely 3-dimensional fecal contamination indicators or a generic tracer, in addition to the circulation variables from the basic OPENCoastS (water levels and 3D velocity, salinity and temperature)  -- The capacity to simulate hindcasts (the past) and scenarios (associated for instance at climate change impacts on coastal systems)  If used in hindcast/scenario mode, simulations do not have to be executed everyday, so requirements on timely delivery of the outputs are alleviated. Like OPENCoastS, the enhanced service uses the efficiently parallelized modeling suite SCHISM. New modules are used, with the associated need to integrate with core eosc-hub services for computing and storage, as OPENCoastS\_PLUS is more demanding of resources than the basic service. |
| **Task** | 5.2 |
| **URL** |  |
| **Service Category** | Data Spaces and Analytics |
| **Service Catalogue** | <https://marketplace.eosc-portal.eu/services/opencoasts-portal> |
| **Location** | Portugal |
| **Duration** | M01-M30 |
| **Modality of access** | Web interfaces |
| **Support offered** | Several training activities are planned |
| **Operational since** | June 2018 |
| **User definition** | Users can be individual researchers or organisations. They can set up deployments for their individual use or for shared purposes (some users will set up deployments while others may just access them). |

### Metrics

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Metric name** | **Baseline** | **Define how measurement is done** | **Period 1 M01-M5** | **Period 2 M06-M10** | **Period 3 M11-M15** |
| No of registered users | 250 | Internal logs of the service | 40 | 10 | 16 |
| No of international deployments | 170 | Internal logs of the service | 6 | 3 | 21 |
| No of requested extensions | 10 | Internal logs of the service | 0 | 7 | 2 |
| No of use cases | 100 | Internal logs of the service | 38 | 9 | 21 |
| No of countries reach | 20 | Internal logs of the service | 26 | 5 | 7 |
| Names of countries reach | Portugal | Internal logs of the service | Australia, Croatia, Netherlands, Taiwan, France, India, Suécia, Brasil, Lithuania, China, USA, Turkey, Jersey, Jamaica, Spain, Iran, Romania, Germany, Martinique, New Zealand, Vietnam, Italy, Indonesia, Ukraine, Senegal, Saudi Arabia, South Korea | Brasil, Portugal, Indonesia, Germany, Chile | Netherlands, Spain, Portugal, Brasil, Austrália, Jordan, Maldives |

### Assessment

A new release of the Data Space (OpenCoastS+) was rolled out into production in February 2022 to assemble on-demand hydrodynamics and water quality forecast systems for selected coastal areas and generate daily 2D and 3D predictions over the region of interest. The new release offers researchers the possibility to predict a vast array of coastal dynamics variables (e.g.: the water levels, wave parameters, 2D and 3D velocities and 3D salinities and temperature), and selected water quality (fecal contamination, generic tracer) indicators in the forecast simulations. It also offers the possibility to compare water inundation extent against processed Sentinel remote sensing images. This new release is now fully integrated with the EGI Check-in[[27]](#footnote-27) service and uses uDocker[[28]](#footnote-28) to load the required software in the HTC and cloud resources, and it is also used as a component for other services (e.g. the EOSC-Synergy WORSICA - for water inundation delimitation). During the second part of the project the OpenCoastS+ Data Space will be further improved integrating additional solution offered by the EOSC Compute Platform, including the Infrastructure Manager (IM)[[29]](#footnote-29) service to deploy and manage the cloud infrastructure, the EGI Data Transfer[[30]](#footnote-30) service to exchange data between cloud and computing clusters, and the EGI Workload Manager[[31]](#footnote-31) service to dispatch jobs across multiple resource providers of the EOSC Compute Platform. The Data Space will also be extended from a forecast-only service to hindcast and forecast runs.

The EOSC Compute Platform resources and the core services are an integral part of the sustainability of the OPENCoastS installation. The availability of these assets is fundamental to open the service to users worldwide in real life applications that have a societal impact either to support management actions under daily (engineering interventions, best periods to access ports, etc.) or emergency (storms, contamination events, etc.) context, or to enhance the capacity of researchers to study in detail processes or the long term impact of climate change or anthropogenic actions.

During the different period of observations we noticed a fluctuation of the metrics reported by the Data Space installation (see table above). This is mainly justified by the development activity undergone during the first part of the project, and the limited opportunities for live dissemination and training events due to COVID restrictions. Even if the new version of OPENCoastS has only been presented once since its release, an overall increase of the VA metrics was reported during the third period of observation. New dissemination activities are planned for the end of 2022 and 2023.

## ENES Data Space

|  |  |
| --- | --- |
| **Description** | ENES data space will deliver a single-entry point to an open and cloud-enabled data science environment for climate data analysis on top of the EOSC Compute Platform implemented in the project. The service brings a data science environment to the end users. It operates on top of the ENES Climate Analytics Service (ECAS), which is one of the EOSC-Hub Thematic Services in EOSC to deliver compute and analytics capabilities to the end users. Compute capacity will be allocated on demand by the EGI-ACE IM/EC3 tool. In addition to that, it will include “synda”, a community tool for data transfer and synchronization, which will be used to set up the climate data archive to be hosted in the ENES data space. Data collections will be shared via EGI DataHub. A JupyterLab front-end will provide the proper (from a data science perspective) entry point to such an environment, which will be enriched with a wide set of open-source scientific Python libraries. The service will provide access to (open) data from the ESGF federated data archive related to large community projects like CMIP6.  The ENES data space will enable analytics capabilities on top of compute (and storage) capacity to support a wide range of data analyses. They include among others: trends, anomaly, climate change signal and extreme events analysis. Single and multi-model experiments will be supported either via interactive (exploratory) or batch data analysis to address different needs and requirements from the end-users. Moreover, the ENES data space is intended to address both data-intensive and data-driven compute scenarios, thus covering a wide spectrum of analytics needs from the community. From an open (data) science perspective, FAIR principles will be pursued; in particular openness and sharing of analytics applications (e.g. Jupyter Notebooks) will be fostered to increase their re-use among users. |
| **Task** | 5.2 |
| **URL** |  |
| **Service Category** | Data Spaces and Analytics |
| **Service Catalogue** | <https://marketplace.eosc-portal.eu/services/enes-climate-analytics-service> |
| **Location** | Hosted in EGI |
| **Duration** | M01-M30 |
| **Modality of access** | Web interfaces |
| **Support offered** | Through a well-defined access workflow for analysis and support, users should gain the required skills to effortlessly use the proposed service.  Support material (training, doc) will be provided in the access portal. General training activities are also foreseen. |
| **Operational since** | M7 (6 months at the beginning of the project will be used to set up the installation in EGI, in particular a preliminary set of relevant data collections, software ecosystem setup, test and validation) |
| **User definition** | Climate scientist/researcher running analytics tasks on the ENES Data Space |

### Metrics

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Metric name** | **Baseline** | **Define how measurement is done** | **Period 1 M01-M05** | **Period 2 M06-M10** | **Period 3 M11-M15** |
| No. of active users | 74 (for the ECAS service) | Internal accounting | N/A | 11 | 24[[32]](#footnote-32) |
| No. of registered users | 40 (for the ECAS service) | Internal accounting | N/A | 16 | 6 |
| No. of use cases | 10 (for the ECASE service) | Internal accounting | N/A | 4 | 2 |
| No. of countries reach | 19 (for the ECAS service) | Internal accounting | N/A | 4 | 5 |
| Names of countries reach | Belgium, Cameroon, Croatia, Denmark, Finland, France, Germany, Greece, India, Italy, Qatar, Romania, Scotland, Slovenia, Spain, Sudan, Turkey, UK, USA | Internal accounting | N/A | Spain, Italy, France, Greece | India, Indonesia, Netherlands, Spain, Italy |

### Assessment

The ENES Data Space[[33]](#footnote-33) is supported by two providers: CMCC and CNRS. This Data Space aims to build an EU data space for large-scale climate data analysis and serve the ENES scientific community. From a technical point of view, the ENES Data Space will deliver a single-entry point to an open and cloud-enabled data science environment for climate data analysis.

The ENES climate modelling community rolled out their new EGI-supported ENES Data Space service into production in November 2021. In the current release, the Data Space provides discovery and interactive analytics on top of a 150TB storage capacity data archive with the use of cloud and HPC resources. The setup is supported by TÜBITAK (TR) and UPV (ES) via dedicated Operational Level Agreements (OLAs)[[34]](#footnote-34). From Q1 2022, periodic new releases of the ENES Data Space are expected every 5 months.

The EOSC Compute Platform is instrumental to support the development of the climate data-science environment. More specifically, EGI-ACE is contributing to scale-up the available computing resources using the Infrastructure Manager (IM)[[35]](#footnote-35) service as IaaS orchestrator and sharing the climate data archives hosted in the ENES data space via the EGI DataHub[[36]](#footnote-36) service. From a technical perspective, the federated computing resources offered by the project help researchers to perform data analysis experiments on large volumes of scientific data (more specifically NetCDF[[37]](#footnote-37) data format for the climate domain) and allow them to address some key challenges and practical issues related to large-scale multi-model data analysis.

In terms of metrics, a negative trend was experienced during the third period of observations compared to the previous one. This is because the Data Space was under development during the first part of the project, and outreach and training activities officially started in Q1 of 2022 when the ENES Data Space was presented in a webinar[[38]](#footnote-38) in March 2022. Moreover, a case study[[39]](#footnote-39) was published for the EGI website to contribute to the promotion and service uptake. On top of this, the ENES Data Space architecture has been significantly enhanced with respect to the first version deployed in the first part of the project, so the service was not always operational during the third reporting period. In the near future, additional training events and dissemination activities are planned (e.g. EGU General Assembly[[40]](#footnote-40), etc.) to further promote the Data Space and reach a wider range of users, thus expecting to increase the values of the collected VA metrics in the next periods of observations.

## PROMINENCE

|  |  |
| --- | --- |
| **Description** | The service will offer model validation based on both experimental and simulated results from any model or experiment where there is a suitable signal based on AI systems. The end goal is to provide a simulation verification service to EOSC users, allowing users to easily run HPC modelling codes and then use AI to compare the simulations to real experimental data. The service will offer access to actual fusion data from the MAST tokamak and possibly EPFL which are currently in the process of opening their data. The service will use the PROMINENCE system that is already available in EOSC and will build on the EGI-ACE HPC and GPU compute resources, OneData for storage, Check-in for AAI, INDIGO DEEPaas for carrying out the ML/DL training and inference, and SimDB for indexing the generated simulation data. It is anticipated that such a service will be of use to any community which performs significant modelling and uses existing experimental data to perform validation. Examples of such communities include astronomy and astrophysics, meteorology, environmental sciences, ecology and biosciences. |
| **Task** | 5.3 |
| **URL** |  |
| **Service Category** | Data Spaces and Analytics |
| **Service Catalogue** | <https://marketplace.eosc-portal.eu/services/prominence> |
| **Location** | UKAEA |
| **Duration** | M01-M30 |
| **Modality of access** | Web interfaces |
| **Support offered** | Technical support for experiments, documentation for the simulation verification service, online tutorials and webinars, training sessions |
| **Operational since** | June 2019 |
| **User definition** | A user from fusion community making use of the Prominence service to produce experimental data, papers, new diagnostic tools and new models describing the behaviour of the plasma  -- Community users will be largely those who have developed the code and already used it in simulations  -- AI experts within the community who are interested in this work and additional AI experts from other institutions |

### Metrics

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Metric name** | **Baseline** | **Define how measurement is done** | **Period 1 M01-M5** | **Period 2 M06-M10** | **Period 3 M11-M15** |
| No of users requesting access | 3 | Internal service monitoring | 0 | 2 | 2 |
| No of jobs submitted | 280 | Internal service monitoring | 371 | 3,122 | 15,495 |
| No of countries reach | 3 | Internal service monitoring | 1 | 1 | 1 |
| Names of countries reach | UK, Korea | Internal service monitoring | UK | UK | UK |

### Assessment

The EOSC Compute Platform is instrumental for the PROMINENCE Data Space[[41]](#footnote-41) to develop a service that will be of use to any community, including Astronomy and Astrophysics, Meteorology, Environmental Sciences, Ecology and Biosciences. Through this PROMINENCE Data Space users can perform significant modelling and use existing experimental data to perform validation. During the first part of the project, a total of 513 vCPU cores, 1.7TB of RAM, 3 GPGPU cards and 60TB of block storage resources were allocated to offer users instant access to the available resources. The capacity allocation is supported by TÜBITAK (TR) and CESGA (ES), UNIV-LILLE (FR) and CESNET-MCC (CZ) with dedicated Operational Level Agreements (OLAs)[[42]](#footnote-42). From a technical perspective, these resources were used by users to easily run HPC modelling codes and use AI models to compare the simulations to real experimental data. The HPC piloting activities to run JOREK non-linear MHD code[[43]](#footnote-43) have started, and the integration with the EGI DataHub[[44]](#footnote-44) service has been completed successfully. The integration of the PROMINENCE Data Space front-end with the EGI Check-in[[45]](#footnote-45) service started in Q4-2021 and planned to be completed in Q2-2022. At the beginning of the project the Data Space planned to use the AI and Machine Learning solutions brought to the EOSC Compute Platform by the DEEP training facility to look for a particular type of disruptive event known as an Edge Localised Mode (ELM) which appear as filaments which move around the outer edge of a plasma. This technical requirement has now a lower priority for this Data Space and the integration with the DEEP tools is not foreseen at the moment.

In terms of metrics, a considerable number of jobs submitted were reported by the Data Space during the third period of observation. This is primarily justified by the use of the computing resources for running plasma simulations and generating datasets. For the second part of the project the PROMINENCE Data Space will continue the integration of the HPC pilot, new documentation will be prepared, and outreach and dissemination activities will be organized to further promote the service uptake.

During the last period of observation, the following % of increment were registered for the Data Space:

* No. of jobs submitted (before the start of the project): 350[[46]](#footnote-46)
* No. of jobs submitted at M15: 15,495, with an increment of +43,27%.

## LOFAR Science Products

|  |  |
| --- | --- |
| **Description** | This service will generate and make available science-ready data, first from LOFAR observational data but to be extended to other existing and future radio astronomical instruments and will provide essential operational experience for a European regional data center for the Square Kilometer Array. Moreover, the generated advanced data products are more easily accessible and usable for cross-domain science and will attract a much wider community than is currently served by the LOFAR Observatory. The service will build on HTC compute infrastructure of the project, and an advanced data product repository and open source pipelines. The service is not operational yet. EOSC-Hub delivered the main software components at the end of 2020, integration into EGI-ACE to be undertaken early 2021. The baseline metrics provided here are for the existing LOFAR archive of observational data. |
| **Task** | 5.3 |
| **URL** |  |
| **Service Category** | Data Spaces and Analytics |
| **Service Catalogue** | <https://marketplace.eosc-portal.eu/services/lofar-science-processing?q=LOFAR+Science+Processing> |
| **Location** | SURFsara |
| **Duration** | M01-M30 |
| **Modality of access** | Web interfaces |
| **Support offered** | The planned activities would encompass user support, training, manuals for the production and access of science ready data and pipelines on HTC infrastructure. |
| **Operational since** | N.A. (the service is to be integrated in 2020 from mature components, and will be outcome of EOSC-hub) |
| **User definition** | Researchers and communities, including RIs:  - Any (radio) astronomer with a need for high quality science-ready images for specific objects or fields observed by LOFAR.  - Advanced data products can be used for multi-frequency research by non-radio astronomers.  - Derived data products are of interest for completely other domains such as ionospheric and space weather research. |

### Metrics

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Metric name** | **Baseline** | **Define how measurement is done** | **Period 1 M01-M5** | **Period 2 M06-M10** | **Period 3 M11-M15** |
| No of researchers requesting access to LOFAR | 0 | Internal service database/accounting | N/A | N/A | 41 |
| Amount of data accessed (PB) | 0 | Internal service database/accounting | N/A | N/A | 0.000096 |
| CPU/hours used for science product generation | 0 | Internal service database/accounting | N/A | N/A | 167 |
| No. of countries reach | 0 | Internal service database/accounting | N/A | N/A | 15 |
| Names of countries reach | - | Internal service database/accounting |  |  | Austria, Belgium, Bulgaria, China, Finland, France, Germany, Ireland, Italy, Latvia, Netherlands, Poland, Ukraine, United Kingdom, United States of America |

### Assessment

The LOFAR Science Processing Data Space[[47]](#footnote-47) provides a service for helping the Radio Astronomy community to generate science-ready LOFAR data, enabling discoveries in astronomy to happen faster and more easily. This Data Space is supported by NWO-I and SURF providers. The first release of the LOFAR Science Processing Data Space was rolled out in production, and on-boarded in the EOSC Portal, at M12. As a direct consequence, initial metrics were collected starting from Period 3. The limited amount of data accessed reported in the table above, is related to a science case that is less demanding in terms of generated data than would typically be expected for LOFAR.

Overall, the resources offered by the EOSC Compute Platform are contributing to open up the processing capabilities of the LOFAR Science Processing Data Space to a wider community of astronomers reducing the technical barriers to get access to high quality science-ready data products that have been observed by LOFAR. From a technical perspective, raw data is processed in order to create a wide range of data products that the LOFAR Science Processing Data Space is offering to the astronomy community.

To further promote the uptake of the LOFAR Science Processing Data Space, domain-specific training events will be organized to promote the Data Space and increase the users base in Europe. In June 2022 the new capabilities offered by this Data Space to the users will be presented during the LOFAR Family meeting. This event is the main gathering for LOFAR users and as such it is a good platform for reaching a wide and relevant audience. In the framework of the EGI-ACE project, the Data Space will also be presented in a webinar in June 2022.

## SeaDataNet WebOcean Data Analysis

|  |  |
| --- | --- |
| **Description** | An online version of the Ocean Data Analysis (ODV) software, which previously was only available as an offline software package. ODV is very popular worldwide among ocean researchers for analysing physical and chemical data collections. WebODV provides interactive exploration, analysis and visualization of oceanographic and other geo-referenced profile or sequence data. SeaDataNet is a successful network and innovator of dedicated data management standards, tools and services, and EGI is a successful partner in providing cloud hosting and computing services for the SeaDataNet research infrastructure. Deploying the online WebODV application at the EGI cloud infrastructure and mobilising its large scientific user basis, consisting of its existing offline users and new users, for trying out and adopting the WebODV cloud version for their science, will provide feedback for further improving the software and its success will have promotional impact on EOSC as an attractive platform for web-based science. |
| **Task** | 5.4 |
| **URL** |  |
| **Service Category** | Data Spaces and Analytics |
| **Service Catalogue** | <https://webodv.awi.de/> |
| **Location** | Alfred Wegener Institute (AWI), Germany |
| **Duration** | M01-M30 |
| **Modality of access** | Web interfaces |
| **Support offered** | User access to the service via registration. Documentation, technical support, training webinars will be provided |
| **Operational since** | 2019 |
| **User definition** | A user is a person in SeaDataNet community or oceanographers from research institutes, universities, and companies. They use the webbodv service to analyse physical and chemical data sets, which are collected at sea using a range of platforms, such as research vessels, small boats, floats, gliders, and others, and a range of instruments such as CTDs, salinographs, water and sediment samplers, ADCPs, and other for studies on marine ecosystems and climate change related analyses |

### Metrics

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Metric name** | **Baseline** | **Define how measurement is done** | **Period 1 M01-M5** | **Period 2 M06-M10** | **Period 3 M11-M15** |
| No of registered users | 100 | SERVICE STILL IN INTEGRATION PHASE | N/A | N/A | N/A |
| No of sessions | 500 | SERVICE STILL IN INTEGRATION PHASE | N/A | N/A | N/A |
| No of countries reach | 35 | SERVICE STILL IN INTEGRATION PHASE | N/A | N/A | N/A |
| Names of countries reach | Mostly European | SERVICE STILL IN INTEGRATION PHASE | N/A | N/A | N/A |

### Assessment

The SeaDataNet WebOcean Data Analysis Data Space is still in development and no metrics were collected so far. A first release of the Data Space is expected during Q2 of 2022. The on-boarding of the Data Space in the EOSC Portal will follow after that.

The SeaDataNet WebOcean Data Analysis Data Space is planned to be accessible by a portal. Through this portal users will be able to access a public instance of the WebOcean service, which will be used to display existing curated datasets from SeaDataNet[[48]](#footnote-48) and ARGO[[49]](#footnote-49) and perform data extraction. During the first part of the project, the public instance has been already installed at INFN-CLOUD-BARI cloud. The portal, which is also integrated with EGI Check-in[[50]](#footnote-50), will also allow users to create private online Ocean Data View (WebODV) instances on the EGI Compute Cloud, where researchers can perform data analysis and have access to a dedicated storage space to use as a workspace. To facilitate the deployment of WebODV private instances transparently to the users, the portal has been integrated also with the EGI PaaS Orchestrator[[51]](#footnote-51).

## EMSO ERIC data services

|  |  |
| --- | --- |
| **Description** | EMSO ERIC data services provide access to harmonized key ocean variables from 11 observatory nodes placed at key environmental sites across European seas, from the North Atlantic, through the Mediterranean, to the Black Sea. The EMSO ERIC data services are currently operated using EGI resources. These services include databases of harmonized EMSO ERIC data and metadata, data portal and dashboards supporting science-driven use case applications, machine-to-machine interfaces, data archive, DAP services, and virtual research environments. |
| **Task** | 5.4 |
| **URL** |  |
| **Service Category** | Data Spaces and Analytics |
| **Service Catalogue** | <http://emso.eu/data/> |
| **Location** | EMSO-ERIC |
| **Duration** | M01-M30 |
| **Modality of access** | Web interfaces |
| **Support offered** | The planned activities would encompass user support, training, and continuous operation of the cloud-enabled systems |
| **Operational since** | 2020 |
| **User definition** | A user is a person making use of EMSO-ERIC data services. |

### Metrics

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Metric name** | **Baseline** | **Define how measurement is done** | **Period 1 M01-M05** | **Period 2 M06-M10** | **Period 3**  **M11-M15** |
| No. of new users | 200 | Internal service database/accounting | 644 | 844 | 489 |
| No. of datasets served | 6,000 | Internal service database/accounting | N/A | N/A | N/A |
| No. of requests served | 15,000 | Internal service database/accounting | 25,414 | 19,255 | 22,653 |
| No. of data analysis run | 200 | Internal service database/accounting | N/A | N/A | N/A |
| No. of countries reach | 30 | Internal service database/accounting | 95 | 90 | 86 |
| Names of countries reach | Worldwide | Internal service database/accounting | Top5: Spain, Italy, China, USA and UK | Top5: China, Italy, Spain, USA, France | Top5: China, Spain, Italy, France, Germany |

### Assessment

EMSO ERIC offers data and services to a large and diverse group of users, from scientists and industries to institutions and policy makers. It is an extraordinary infrastructure to provide relevant information for defining environmental policies based on scientific data. The EOSC Compute Platform is contributing to operate and deliver the EMSO ERIC data services[[52]](#footnote-52) and provide access to scientific data.

The resources and the solutions provided by the EOSC Compute Platform are instrumental to support the transition of the EMSO-ERIC data services from the pre-production to the production level. To support this transition, and maintain the operation of the EMSO ERIC data services, the project is contributing providing access to the federated resources of the EOSC Compute Platform and offering Authorization and Authentication solution (via EGI Check-in) to enable federated access of users to the EMSO ERIC data services. For what concerns the federated resources, a total of 492 vCPU cores, 1.7TB of RAM and 10.6TB of block storage resources have been allocated with a dedicated Service Level Agreement (SLA)[[53]](#footnote-53) which has been extended until 06/2023.

In terms of metrics, as reported in the different periods of observation (see table above), most of them already met the target baseline. More specifically, a progressively increasing number of requests served by the EMSO ERIC Data Portal was registered over the three periods of observations. The two metrics: ‘No. of datasets served’ and ‘# of data analysis run’ were not provided as they are no longer representative to describe the level of maturity, nor explain how the Data Space is performing. For this reason, a request to remove/update these two metrics has been already submitted for the next project amendment.

During the last period of observation, the following % of increment were registered by the Data Space:

* No. of new users (before the start of the project): 250[[54]](#footnote-54)
* No. of new users at M15: (644+844+489)=1977, with an increment of +690%. If we analyse the metric in the last 2 periods of observations, we notice a drop. This was primarily due to a significant reduction of use of the EMSO ERIC data services during the Christmas break and the lack of any important event during the last reporting period.
* No. of requests served (before the start of the project): 18,75048
* No. of requests served at M15: 22,653, with an increment of +20.81%.

## GBIF Cloud data space

|  |  |
| --- | --- |
| **Description** | In the framework of EOSC-hub, a number of services of GBIF Spain have already been made available through the EOSC Portal. This installation will be an integrated platform hosting all the GBIF data from all Iberian GBIF publishers, plus data from other GBIF publishers for the Iberian region and integrated storage and data analytics capabilities to support researchers perform data processing and visualisation online. The service integrated biodiversity data and geospatial data (climate, soil, land use, environmental variables, etc.). The service will provide new facilities in 3 areas: Localization (serving national conservation strategies, EU directives); More advanced visualisation and analysis capabilities; Integration between biodiversity data and geospatial data. The platform will orchestrate all the necessary EOSC/EGI services (AAI, EGI Jyputer notebook), made available through the EOSC Portal, and will support researchers in the area of biodiversity by connecting those data, with a layer of advanced computing and storage services. |
| **Task** | 5.4 |
| **URL** |  |
| **Service Category** | Data Spaces and Analytics |
| **Service Catalogue** | <https://marketplace.eosc-portal.eu/services/e-learning-platform-of-gbif-spain>  <https://marketplace.eosc-portal.eu/services/gbif-spain-occurrence-records>  <https://marketplace.eosc-portal.eu/services/gbif-spain-collections-registry>  <https://marketplace.eosc-portal.eu/services/gbif-spain-images-portal>  <https://marketplace.eosc-portal.eu/services/gbif-spain-regions-module>  <https://marketplace.eosc-portal.eu/services/gbif-spain-spatial-portal>  <https://marketplace.eosc-portal.eu/services/gbif-spain-species-portal>  <https://marketplace.eosc-portal.eu/services/gbif-portugal-occurrence-records?q=GBIF+Portugal+Occurrence+Records> |
| **Location** | Madrid and Lisbon |
| **Duration** | M01-M30 |
| **Modality of access** | Web interfaces |
| **Support offered** | It is planned to organise training events and workshops as needed. A set of manuals and video tutorials will be available for some of the services. |
| **Operational since** | Services are already available in datos.gbif.es since 2015 |
| **User definition** | A user of GBIF service, who can be:  -- Researchers in biodiversity, and ecosystems in global change  -- Policy making in the framework of the EU Green Deal directives  -- Companies and the public sector working on land conservation and restoration, invasive species management, cultural ecosystem services |

### Metrics

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Metric name** | **Baseline** | **Define how measurement is done** | **Period 1 M01-M5** | **Period 2 M06-M10** | **Period 3 M11-M15** |
| No of unique users | 8,031 | Internal logs | 2,721 | 1,472 | 2,665 |
| No of sessions opened | 13,605 | Google Analytics | 4,635 | 2,712 | 5,457 |
| No of page views | 79,071 | Google Analytics | 30,717 | 18,802 | 45,643 |
| Data stored (TB) | 34 | Internal logs | 6.6 | 7.3 | 10.6 |
| No of downloads | 42,363 | Internal logs | 652 | 249 | 672 |
| No of registries downloads | 426,795,567 | Internal logs | 652 | 14,456,047 | 42,502,362 |
| No of registries visualized | 51,907,016 | Internal logs | 2,384,996 | 2,443,061 | 2,466,553 |
| No of countries reach | 59 | Google Analytics | 115 | 42 | 56 |
| Names of countries reach | Worldwide | Google Analytics | Spain 60 % , US 4.9 %, China 3.6 %, Ecuador 2 %, Germany 1.6 %, Brazil 1.6 %, Mexico 1.5 %, France 1.25 %, Colombia 1.1 %, Others 21.1 % | Spain 70.1%, USA 10 %, China 5.8 %, Colombia 1.8 %, Mexico 1.8 %, Argentina 0.8 %,Germany 0.8 %, Finland, 0.7 %, UK 0.7 %,Others 7.5 % | Spain 78.8%, USA 5.6%, China 2.4%, Mexico 1.5%, Others 11.7% |

### Assessment

GBIF (the Global Biodiversity Information Facility)[[55]](#footnote-55) is an international network and data infrastructure funded by the world's governments aimed at providing anyone, anywhere, open access to biodiversity data. In the context of the EGI-ACE project the Data Space is supported by the two providers: LIP and CSIC. During the first part of the project the GBIF Cloud Data Space focused on the integration of datasets hosted from GBIF Spain and Portugal. This integration activity required primarily architectural support from the project. As a result of this integration, the GBIF Portugal Occurrence Records[[56]](#footnote-56) was registered in the EOSC Portal in Feb. 2022. As a next step, the GBIF Cloud Data Space is now validating the solution with real data before publishing the integrated portal in EOSC. This is expected for Q3 of 2022. For the second part of the project the current AAI model, based on CAS Enterprise Single Sign-On, will be further extended with the federated authentication and authorization framework supported by the EGI-ACE project. Additionally, the benefits of using EGI Jupyter notebooks as an additional interface to access data for further exploitation will be investigated.

The collaboration started in the framework of the EGI-ACE project is strategic for the deployment in EOSC of a European platform to support the integration of biodiversity data. Through this platform the biodiversity data, collected from different European countries, will be offered to end-users without potential breaks that political borders may impose on data.

In terms of metrics, fluctuations were reported during the second period of observation (see table above). These fluctuations are primarily justified by the integration activity undergone during the first part of the project. With the validation of the architecture, we expect these fluctuations to disappear and increase the values of the metrics collected.

During the last period of observation, the following % of variation were registered by the Data Space:

* No. of unique users (before the start of the project): 10,038[[57]](#footnote-57)
* No. of unique users at M15 (normalized): (2,721+1,472+2,665)=6,858. Compared to the baseline, the % of variation is still negative (-31,68%) however, a significant increase was reported compared to the second period of observation. The same analysis applies to other metrics reported in the table above.

## Disaster mitigation and agriculture

|  |  |
| --- | --- |
| **Description** | As a result of EGI-Engage and EOSC-hub projects the disaster mitigation communities from the Asia-Pacific region developed and offered 2 simulation portals in EOSC (tsunami wave propagation simulations and for WRF-based weather simulation). These installations will continue and will be expanded with three simulation portals for fire/haze/smoke monitoring, flood, typhoon/cyclone, tsunami, storm surge and agriculture research - based on Asia-Pacific and European expertise, datasets and regional e-infrastructures. The user community has been extended from 6 core countries (TW, PH, VN, MY, TH, ID) to open participation via APAN, including agriculture, remote sensing, and biodiversity & ecological monitoring, also teamed up with Sentinel Asia and earth observation communities to support the quantitative hazard risk analysis and disaster management in Asia and Europe. |
| **Task** | 5.4 |
| **URL** |  |
| **Service Category** | Data Spaces and Analytics |
| **Service Catalogue** |  |
| **Location** | Taiwan, AS(ASGC) |
| **Duration** | M01-M30 |
| **Modality of access** | Web interfaces |
| **Support offered** | Helpdesk, Technical support, Training and Webinars |
| **Operational since** | 2018 |
| **User definition** | Researchers working in disaster simulation, agricultural monitoring, land observation, civil protection. |

### Metrics

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Metric name** | **Baseline** | **Define how measurement is done** | **Period 1 M01-M5** | **Period 2 M06-M10** | **Period 3 M11-M15** |
| No. of new registered users | 150 | Internal service database/accounting | N/A | N/A | N/A |
| No. of simulation jobs run | 800 | Internal service database/accounting | N/A | N/A | N/A |
| No. of countries reach | 13 | Internal service database/accounting | N/A | N/A | N/A |
| Names of countries reach | Taiwan, Philippines, Vietnam, Thailand, Myanmar, India, Japan, Malaysia, Bangladesh, Indonesia, Czech Republic, Italy, Germany | Internal service database/accounting | N/A | N/A | N/A |

### Assessment

The Disaster Mitigation and Agriculture is one of the unfunded Data Space supported by the project. To demonstrate the integration of advanced distributed cloud infrastructure and simulation portals, case studies as well as knowledge discovery in the Asia region for reduction of natural hazard impacts in various aspects, the Data Space focuses on the following tasks:

* Develop and re-organize simulation portals to make use of EOSC cloud in Asia and Europe on hazards including weather, fire/haze/smoke monitoring, flood, typhoon/cyclone, tsunami, and storm surge.
* Conducting a case study in Thailand to understand weather parameters impact on agriculture production (JP, TH, PH, TW).
* Keep on supporting case studies proposed by partners in Asia and improving the risk analysis workflow and functionality of simulation portals coordinated by Taiwan (ASGC), such as flood (MM, VN, MY), lightning (BD), storm surge and tsunami (PH, ID), forest fire/haze/maze monitoring (TH and ASEAN).

In the context of the EGI-ACE project:

* The iCOMCOT portal recently passed the validation by scientists in 2021. The iCOMCOT portal is currently running in the Asian cloud infrastructure composed of ~1000 vCPU cores, 8 GPU cards, and 20 TB of storage. To encourage resource sharing and offer researchers access to the EOSC Compute Platform to run numerical models and improve the risk analysis workflow and functionality of the simulation portals, the vo.environmental.egi.eu[[58]](#footnote-58) Virtual Organization (VO) was configured in the EGI Operations Portal and initial resources of 16 vCPU cores, 48 GB of RAM, and 1 TB of block storage were allocated in one of the providers of the EGI Federation. During the second part of the project the iCOMCOT portal will be connected to the EOSC Compute Platform. Additionally, the benefits of using EGI Jupyter notebooks as an additional interface to access data for further exploitation will be also investigated. The on-boarding of the iCOMCOT portal by the service provider in the EOSC Portal is expected for Q4 of 2022.
* The Storm Surge portal is currently in validation, and it is expected to be available in Q4 of 2022.

In terms of dissemination and outreach activities, several events were organized in the Asian region to promote the collaboration around national hazards in the Asia region, including the APAN53[[59]](#footnote-59) and the ISGC 2022[[60]](#footnote-60) conferences.

## OPERAS Metrics service and Certification service

|  |  |
| --- | --- |
| **Description** | Developed in HIRMEOS. The service collects usage and impact metrics related to Open Access monographs from various different sources and allows for their access, display and analysis from a single access point. The OPERAS Metrics Suite consists of a shared data model, various open-source tools and services designed to serve the various components used by the shared OPERAS database and API used for a diverse range of usage and impact metrics including downloads, web visits, tweets, Wikipedia mentions, etc. The Certification service is operated by DOAB, which collects the variety of peer reviewing practices from hundreds of monograph publishing houses, categorizes them, and provides a single access point to the list of certified peer reviewed monographs available in Open Access in the world. DOAB is a digital directory of peer-reviewed Open Access books and Open Access book publishers. The primary aim of the service is to increase discoverability of OA books so that they can reach a broader audience. The Certification Service operates as a quality insurance service for the benefit of readers and the service providers working with them, such as the libraries. |
| **Task** | 5.5 |
| **URL** |  |
| **Service Category** | Data Spaces and Analytics |
| **Service Catalogue** | <https://metrics.operas-eu.org/> |
| **Location** | Metrics: Hosted in EGI (site to be defined at the beginning of the project), Certification: Hosted by Huma-Num at IN2P3 (France) |
| **Duration** | M01-M30 |
| **Modality of access** | Web interfaces and APIs |
| **Support offered** | The planned activities would encompass user support, training, and continuous operation of the service and core EGI services (e.g. AAI). The Metrics service will use the EGI Cloud Container Compute (Kubernetes). Technical and scientific staff directly working for the provision of virtual access (Ubiquity Press for OPERAS AISBL). |
| **Operational since** | June 2019 |
| **User definition** | A typical user is publishers, SSH researchers, research engineers, librarians, also citizens |

### Metrics

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Metric name** | **Baseline** | **Define how measurement is done** | **Period 1 M01-M5** | **Period 2 M06-M10** | **Period 3 M11-M15** |
| No of registered publishers | 6 (Metrics) + 7 (Certification) | Internal | Same as baseline | N/A | N/A |
| No of API hits per day | 100,000 | Internal | Same as baseline | N/A | N/A |
| No of countries reach | 237 (Metrics) + 5 (Certification) | Internal | Same as baseline | N/A | N/A |
| Names of countries reach | France, UK, Sweden, Germany, Greece | Internal | Same as baseline | N/A | N/A |

### Assessment

The OPERAS Metrics service[[61]](#footnote-61) is one of the unfunded Data Space supported by the project. An initial resource capacity allocation was activated at IN2P3-IRES in France with 10 vCPU cores, 20GB of RAM and 1TB of block storage[[62]](#footnote-62). It is important to note that the planned usage of the EGI infrastructure, supported by the EGI-ACE project, is to support an operational level service of OPERAS. Therefore, this means that a carefully planned technical shift from the current operational environment into a new environment within the EGI infrastructure required a higher-than-average time and effort. Though important for the long-term, due to the lack of staff, the workplan was delayed regarding the planned usage of the resources provided, which is the main reason why no metrics were reported in the three periods of observation.

In addition to the resources allocated as part of EGI-ACE, IN2P3-IRES has pledged continued support beyond the life of the project that would be re-evaluated on an annual basis, which was an important consideration with regards to the investment of not only the time required to make the shift of technical environments but provide the confidence in longer-term stability of the operational service. This agreement has only recently been discussed, but the subsequent work has already begun between the OPERAS metrics development teams and the EGI-AGE technical support to facilitate the deployment and the operation, which is based on docker containers that are deployed in a Kubernetes cluster, and the support to the Infrastructure Manager (IM) was enabled. Instructions to deploy and operate Kubernetes clusters in the EOSC Compute Platform with the IM were shared with the main contacts therefore the workplan is in progress.

This work is expected to be completed in the next 6 months with the first release of the OPERAS Metrics service installation. During the second part of the project, the OPERAS Metrics service installation will be further integrated in the EOSC Compute Platform.

# 

# Satisfaction

In this section we report the customers’ satisfaction of the Data Space installations. In particular, those that are operating at pre-production/production level are taken into consideration.

## The WeNMR Thematic Services

The WeNMR Thematic Services by design include a mechanism to constantly monitor the level of satisfaction of the services offered to their users. Customers' feedback is mainly used for improving the performance and the functionalities offered by the services. During the reporting period, 12 training events were organized by WeNMR. Overall, the level of satisfaction[[63]](#footnote-63) received by the WeNMR Thematic portal is shown in the Table 6.

Table 6 - WeNMR Thematic Services satisfaction (source <https://wenmr.science.uu.nl/stats>)

|  |  |
| --- | --- |
| **Thematic Services** | **User’s feedback** |
| **DisVis Portal** | 4.8 (from 77 respondents) |
| **HADDOCK2.4** | 4.9 (from 1,652 respondents) |
| **PowerFit Portal** | 4.8 (from 19 respondents) |
| **SpotOn Portal** | 4.7 (from 82 respondents) |

## The Virtual Imaging Platform

As part of the webinar programme organized by T2.3, the status of the Virtual Imaging Platform (VIP) Data Space installation was introduced in 2022. The VIP webinar was attended by 46 participants from 25 different countries. The overall feedback received during the webinar was positive as shown in the Table 7:

Table 7 - Customers’ satisfaction feedback of the VIP Data Space

|  |  |  |
| --- | --- | --- |
| **Webinar Title** | **Overall Webinar[[64]](#footnote-64)** | **Content of the Webinar** |
| The Virtual Imaging Platform: Scientific Applications as a Service and Beyond[[65]](#footnote-65) | 4.2 (from 9 respondents) | 4.3 (from 9 participants) |

## The ENES Data Space

Also the ENES Data Space was introduced in a webinar in 2022 as part of the webinar series organized by T2.3. The webinar was attended by 44 participants from 18 different countries. The overall feedback received from participants is shown in Table 8:

Table 8 - Customers’ satisfaction feedback of the ENES Data Space

|  |  |  |
| --- | --- | --- |
| **Webinar Title** | **Overall Webinar60** | **Content of the Webinar** |
| The ENES Data Space Service[[66]](#footnote-66) | 4.5 (from 13 respondents) | 4.5 (from 13 respondents) |

## EMSO ERIC Data Service

On 20-22 October 2021, the resources of the EOSC Compute Platform were used to host the EMSO Time Series training[[67]](#footnote-67) session that took place in the Canary Islands, Spain. During the event, in particular during the hands-on sessions, participants explored how to profit from the EOSC computing infrastructure using Jupyter notebooks to provide cost-efficient information for the assessment of marine mammal populations, the detection of fish reproduction areas, the detection of greenhouse gas seeps from pipelines and deep-sea carbon storage, gasification of methane clathrates, adverse meteorological conditions, detection of low-frequency seismic events, ice-cracking, ocean basin sound-velocity tomography and acoustic communication.

# Service Orders

For the Data Space installations already registered in the EOSC Portal Catalogue and Marketplace[[68]](#footnote-68), we report here the statistics of the service orders received during the first 15 months of the project. These statistics were collected from the EOSC Metric Portal[[69]](#footnote-69).

|  |  |
| --- | --- |
| **WP5 Data Space installations** | **Service Orders** |
| **Haddock2.4 Web Portal** | 3 |
| **DisVis Portal** | 2 |
| **European Galaxy Server** | 2 |
| **AMBER-Based Portal Server** | 1 |
| **Powerfit Web Portal** | 1 |
| **ENES Climate Analytics Service** | 1 |
| **OpenCoastS Portal** | 1 |

Table 9 - Number of Service Orders (SOs) related to WP5 Data Space installations (during period Jan. 2021 - March 2022)

**Appendix I - Status of the WP5 integration activities**

Figure 10 - Service adoption within the Data Space Installations

**Graphical user interface, application

Description automatically generated**

1. Including number of users of the <https://operations-portal.egi.eu/vo/view/voname/vo.enes.org> VO [↑](#footnote-ref-1)
2. It also includes the users in the <https://operations-portal.egi.eu/vo/view/voname/vo.enes.org> VO [↑](#footnote-ref-2)
3. <https://www.uu.nl/en> [↑](#footnote-ref-3)
4. <https://marketplace.eosc-portal.eu/services/c/software?scientific_domains%5B%5D=23> [↑](#footnote-ref-4)
5. <https://documents.egi.eu/document/2751> [↑](#footnote-ref-5)
6. Normalized metric (considering the last 15 months) [↑](#footnote-ref-6)
7. <https://www.cerm.unifi.it/about-us/cirmmp> [↑](#footnote-ref-7)
8. <https://www.egi.eu/use-cases/research-infrastructures/wenmr-a-worldwide-e-infrastructure-for-nmr-60348-2/> [↑](#footnote-ref-8)
9. <https://www.egi.eu/services/workload-manager/> [↑](#footnote-ref-9)
10. <https://marketplace.eosc-portal.eu/services/virtual-imaging-platform> [↑](#footnote-ref-10)
11. <https://www.egi.eu/services/check-in/> [↑](#footnote-ref-11)
12. <https://vip.creatis.insa-lyon.fr/> [↑](#footnote-ref-12)
13. <https://indico4.twgrid.org/event/20/contributions/1117> [↑](#footnote-ref-13)
14. <https://cbica.github.io/CaPTk/preprocessing_brats.html> [↑](#footnote-ref-14)
15. <https://portal.fli-iam.irisa.fr/msseg-2/> [↑](#footnote-ref-15)
16. https://indico.egi.eu/event/5464/contributions/15641/attachments/14127/17960/EGI\_VIP\_MSSEG2\_v3.pdf [↑](#footnote-ref-16)
17. <https://documents.egi.eu/document/2874> [↑](#footnote-ref-17)
18. <https://www.egi.eu/use-cases/research-infrastructures/biomed/> [↑](#footnote-ref-18)
19. <https://indico.egi.eu/event/5824/> [↑](#footnote-ref-19)
20. Normalized metric (considering the last 15 months) [↑](#footnote-ref-20)
21. <https://documents.egi.eu/document/3749> [↑](#footnote-ref-21)
22. <https://marketplace.eosc-portal.eu/services/european-galaxy-server> [↑](#footnote-ref-22)
23. <https://www.egi.eu/services/check-in/> [↑](#footnote-ref-23)
24. <https://www.egi.eu/services/datahub/> [↑](#footnote-ref-24)
25. <https://www.egi.eu/services/workload-manager/> [↑](#footnote-ref-25)
26. Normalized metric (considering the last 15 months) [↑](#footnote-ref-26)
27. <https://www.egi.eu/services/check-in/> [↑](#footnote-ref-27)
28. <https://github.com/indigo-dc/udocker> [↑](#footnote-ref-28)
29. <https://www.grycap.upv.es/im/index.php> [↑](#footnote-ref-29)
30. <https://www.egi.eu/services/data-transfer/> [↑](#footnote-ref-30)
31. <https://www.egi.eu/services/workload-manager/> [↑](#footnote-ref-31)
32. Including number of users of the <https://operations-portal.egi.eu/vo/view/voname/vo.enes.org> VO [↑](#footnote-ref-32)
33. <https://marketplace.eosc-portal.eu/services/enes-data-space> [↑](#footnote-ref-33)
34. <https://documents.egi.eu/document/3835> [↑](#footnote-ref-34)
35. <https://marketplace.eosc-portal.eu/services/infrastructure-manager-im> [↑](#footnote-ref-35)
36. [https://marketplace.eosc-portal.eu/services/egi-datahub](https://marketplace.eosc-portal.eu/services/egi-datahub?q=EGI+DataHub) [↑](#footnote-ref-36)
37. <https://www.unidata.ucar.edu/software/netcdf/> [↑](#footnote-ref-37)
38. <https://indico.egi.eu/event/5743/> [↑](#footnote-ref-38)
39. <https://www.egi.eu/use-cases/research-infrastructures/enes> [↑](#footnote-ref-39)
40. <https://www.egu.eu/meetings/general-assembly/> [↑](#footnote-ref-40)
41. <https://marketplace.eosc-portal.eu/services/prominence> [↑](#footnote-ref-41)
42. <https://documents.egi.eu/document/3484> [↑](#footnote-ref-42)
43. <https://www.jorek.eu/> [↑](#footnote-ref-43)
44. <https://www.egi.eu/services/datahub/> [↑](#footnote-ref-44)
45. <https://www.egi.eu/services/check-in/> [↑](#footnote-ref-45)
46. Normalized metric (considering the last 15 months) [↑](#footnote-ref-46)
47. <https://marketplace.eosc-portal.eu/services/lofar-science-processing> [↑](#footnote-ref-47)
48. <https://www.seadatanet.org/> [↑](#footnote-ref-48)
49. <https://monitoring.seadatanet.org/sdc/dashboard/> [↑](#footnote-ref-49)
50. <https://www.egi.eu/services/check-in/> [↑](#footnote-ref-50)
51. <https://marketplace.eosc-portal.eu/services/paas-orchestrator?q=PaaS+Orchestrator> [↑](#footnote-ref-51)
52. <https://marketplace.eosc-portal.eu/services/emso-eric-data-portal> [↑](#footnote-ref-52)
53. <https://documents.egi.eu/document/3539> [↑](#footnote-ref-53)
54. Normalized metric (considering the last 15 months) [↑](#footnote-ref-54)
55. <https://www.gbif.org/> [↑](#footnote-ref-55)
56. <https://marketplace.eosc-portal.eu/services/gbif-portugal-occurrence-records> [↑](#footnote-ref-56)
57. Normalized metric (considering the last 15 months) [↑](#footnote-ref-57)
58. <https://operations-portal.egi.eu/vo/view/voname/vo.environmental.egi.eu> [↑](#footnote-ref-58)
59. <https://apan53.apan.net/> [↑](#footnote-ref-59)
60. <https://indico4.twgrid.org/event/20/> [↑](#footnote-ref-60)
61. <https://marketplace.eosc-portal.eu/services/operas-metrics-service> [↑](#footnote-ref-61)
62. <https://documents.egi.eu/document/3712> [↑](#footnote-ref-62)
63. The level of satisfaction is measured from 1 (min) to 5 (max). [↑](#footnote-ref-63)
64. On a scale of 1 to 5 with 5 being highest. [↑](#footnote-ref-64)
65. <https://indico.egi.eu/event/5824/> [↑](#footnote-ref-65)
66. <https://indico.egi.eu/event/5743/> [↑](#footnote-ref-66)
67. <https://tsc2021.emso.eu/> [↑](#footnote-ref-67)
68. <https://marketplace.eosc-portal.eu/> [↑](#footnote-ref-68)
69. [https://opsportal.eosc-portal.eu/metricsEOSC/ServiceOrder/2021-01-01/2022-03-31/stats/on#](https://opsportal.eosc-portal.eu/metricsEOSC/ServiceOrder/2021-01-01/2022-03-31/stats/on) [↑](#footnote-ref-69)