KER 3: Research data spaces and processing tools for EOSC

Final Review meeting, 12/09/2023

Giuseppe La Rocca
EGI-ACE WP5 Leader
Community Support Team Lead. at EGI Foundation

Dissemination level: Public

Disclosing Party: Project Consortium

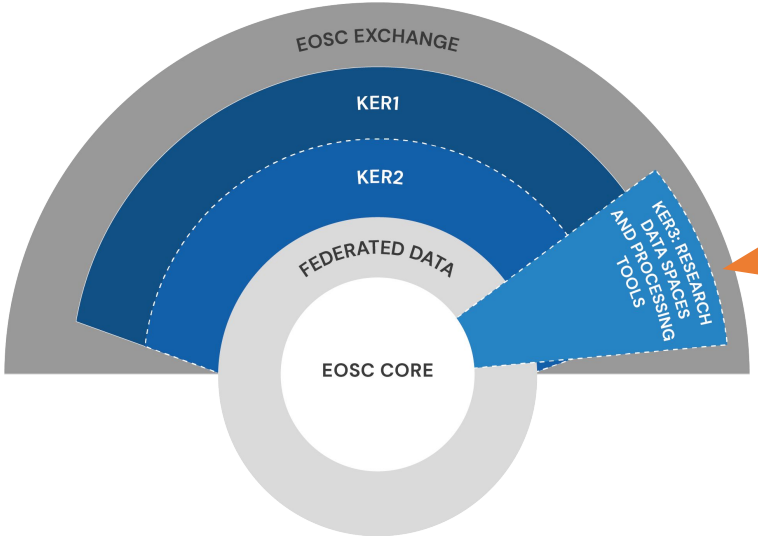
Recipient Party: European Commission

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

Outline

- KER3 overview
- Status of the EGI-ACE Thematic Services at M30
- Dissemination and Outreach activities
- KPI
- Future outlook
- KER3 Demo

KER OVERVIEW

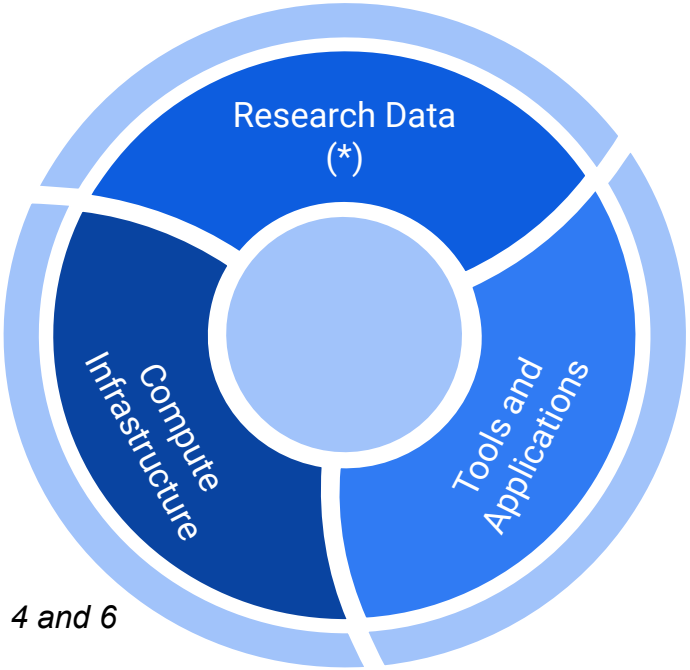


KER3: RESEARCH DATA SPACES AND PROCESSING TOOLS FOR EOSC

- Health and Medicine
- Social Sciences and Humanities
- Climate Research
- Energy and Physical Sciences
- Environmental Sciences
- New Thematic Services

- **Provided by:** WP5, T2.3

What we mean by EGI-ACE Data Space?



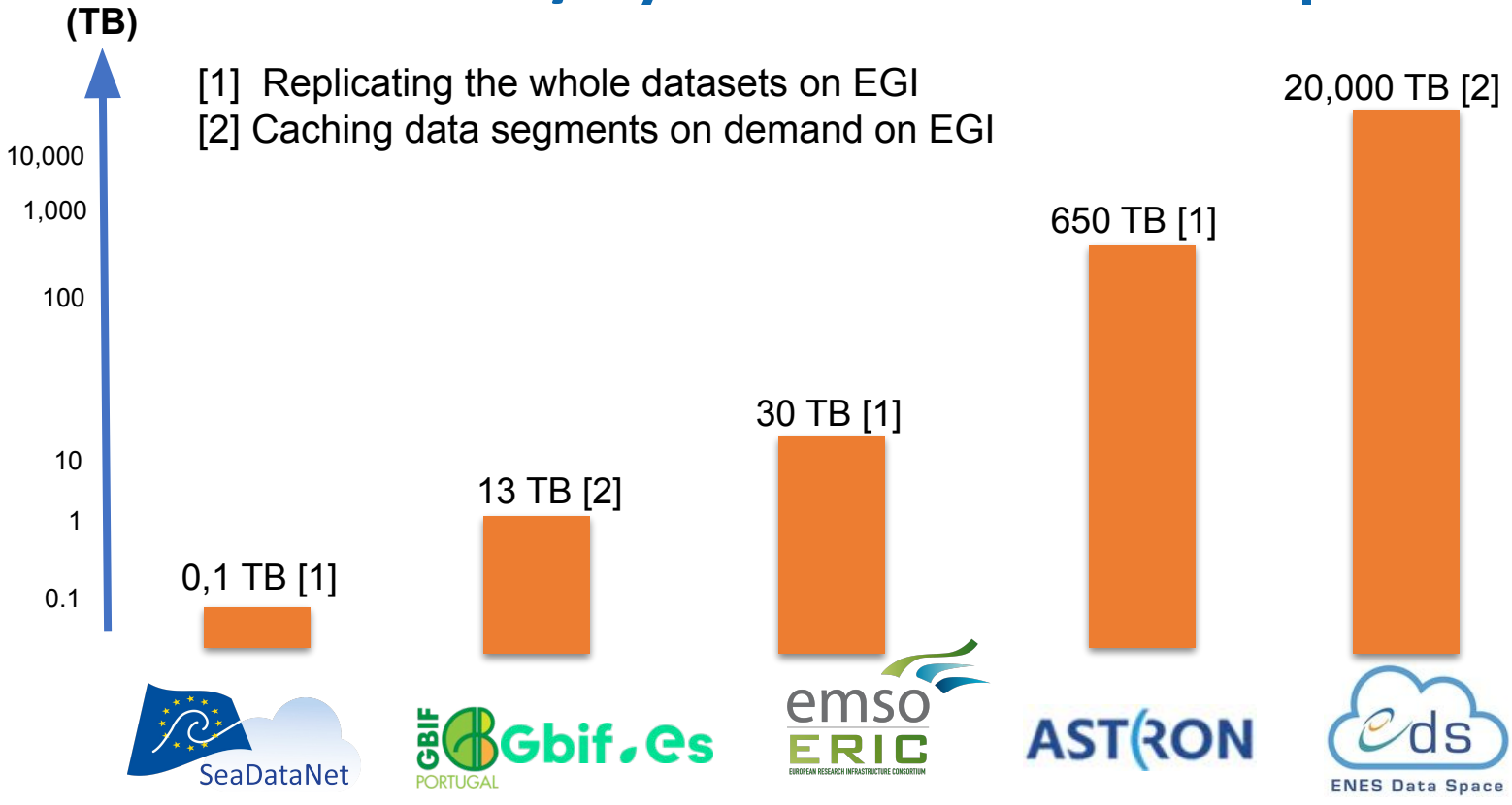
WP3, 4 and 6

- EGI-ACE provided **solutions** for the exploitation of big, open scientific data.
- WP5 provided effort for integration, delivery, training, support (**)
- EGI's **policies** ensure that Data Spaces are established and managed in a way that promotes trust, transparency, and accountability.
- Supported by federated resources, access and platform services with **Service Level Agreements** (SLAs)

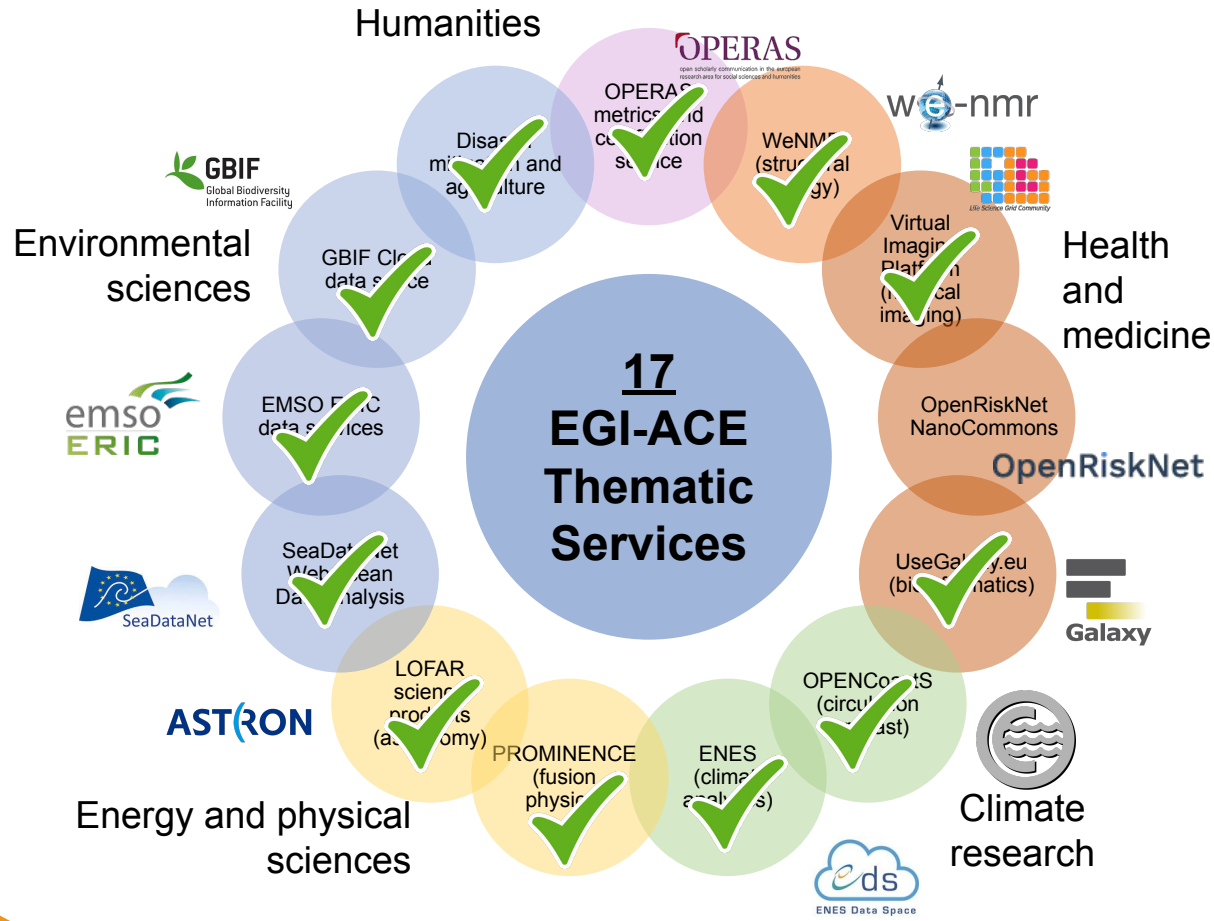
(*) Datasets are curated and made available by research organizations outside EGI-ACE
(**) Tools and Applications are also developed outside EGI-ACE

Data delivery by the EGI-ACE Data Spaces

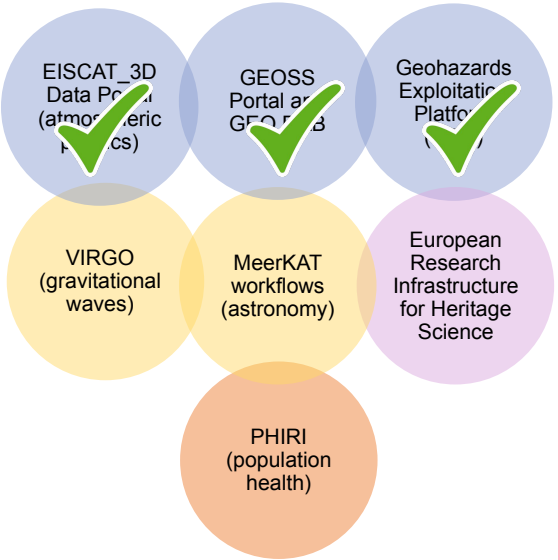
[1] Replicating the whole datasets on EGI
[2] Caching data segments on demand on EGI



Status of the Thematic Services at M30



7 Early Adopters



Already on-boarded in the EOSC Portal

WP3,4, 6 services integration map

Table legend

Adopted	Technology/service already integrated before EGI-ACE
Planning	Integration is ongoing
Integrated	Technology/service was integrated in the context of EGI-ACE
Investigating	Technology/service is considered for adoption, but the integration and assessment work is yet to start

	WeNMR services suite	Virtual Imaging Platform	OpenRiskNet	useGalaxy.eu	OPENCoastS	ENES Data Space	PROMINENCE	LOFAR Science Processing	SeaDataNet WebOcean Data Analysis	EMSO-ERIC	GBIF Cloud Data Space	Disaster Mitigation and Agriculture	OPERAS Metrics and PRISM serivces
WP3, WP4, WP6													
EGI Cloud Compute	Adopted	Adopted	Adopted		Adopted	Integrated	Adopted		Integrated	Adopted	Planning	Planning	Adopted
EGI High-Throughput Compute	Adopted	Adopted			Adopted			Integrated					
EGI Online Storage	Adopted	Adopted	Adopted		Adopted	Integrated	Adopted	Integrated	Integrated	Adopted	Planning	Planning	Adopted
Infrastructure Manager					Planning	Integrated							
EGI Workload Manager	Adopted	Adopted		Investigating	Adopted								
EGI Notebooks		Investigating									Investigating	Investigating	
CVMFS	Adopted								Integrated				
EGI Check-in	Adopted	Integrated	Investigating	Investigating	Integrated	Integrated	Integrated		Integrated	Adopted			Adopted
EGI DataHub				Investigating	Planning	Planning	Integrated	Investigating					
PaaS Orchestrator									Integrated				

	EISCAT_3D	VIRGO	e-RIHS	GEO-DAB	TEPs	PHIRI	MerKaaT
WP3, WP4, WP6							
EGI Cloud Compute	Adopted	Integrated	Adopted	Adopted	Adopted	Integrated	
EGI High-Throughput Compute							
EGI Online Storage	Adopted	Integrated	Adopted	Adopted	Adopted	Integrated	
Infrastructure Manager							
EGI Workload Manager	Adopted						
EGI Notebooks	Integrated					Investigating	
CVMFS	Adopted						
EGI Check-in	Integrated		Investigating	Adopted	Planning	Investigating	
EGI DataHub							
PaaS Orchestrator		Integrated					

Zoom in one of the EGI-ACE Data Spaces

European Network for Earth System Modelling

The European Network for Earth System modelling (ENES) network, was launched in 2001

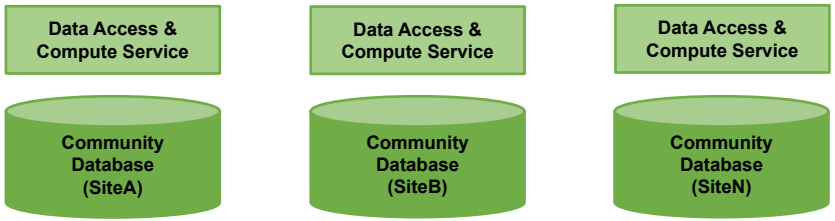
Accelerate progress in Earth's climate system modelling

- Data processing,
- Archiving,
- Sharing, and
- Analysis.

enes
EUROPEAN NETWORK
FOR EARTH SYSTEM MODELLING

The ENES Data Space architecture

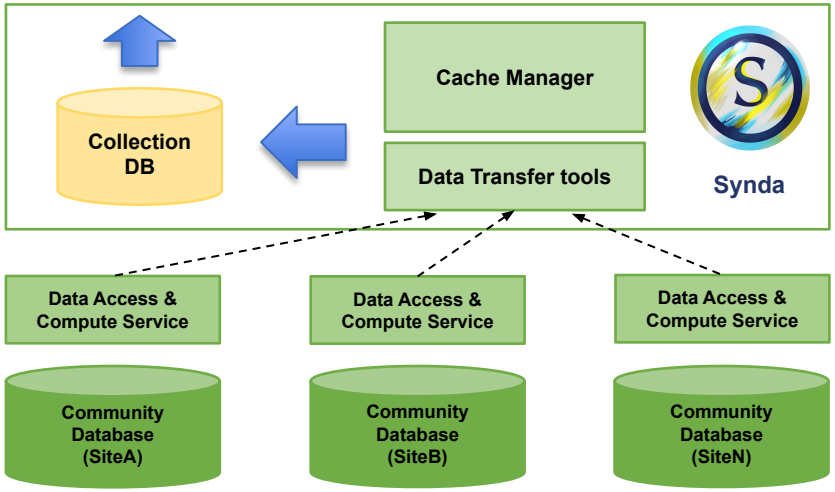
The ENES Data Space architecture



Community (legacy) Infrastructure
Community-specific data access services & interfaces

is-enes http://is.enes.org/ ESGF Earth System Grid Federation

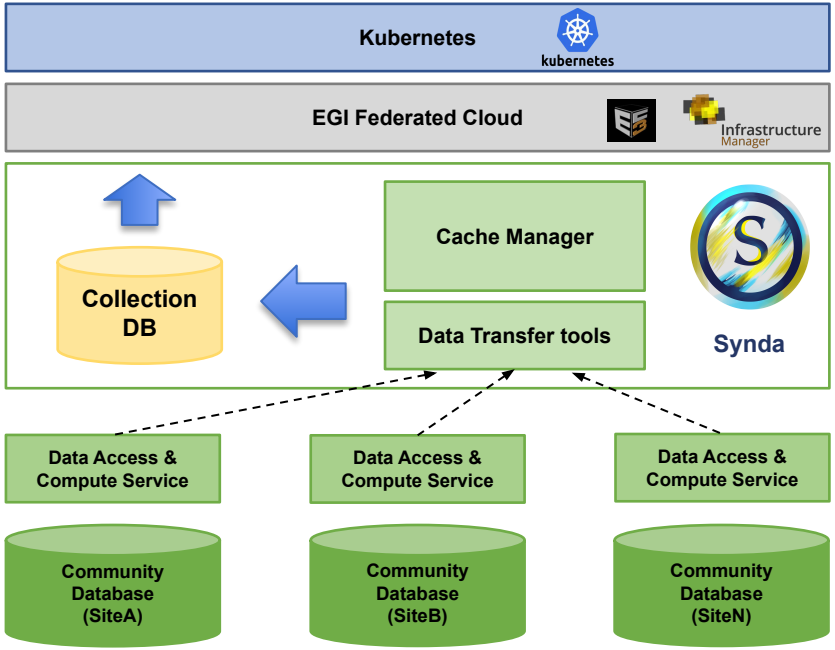
The ENES Data Space architecture



Data collector and Cache Service

Community (legacy) Infrastructure
Community-specific data access services & interfaces

The ENES Data Space architecture



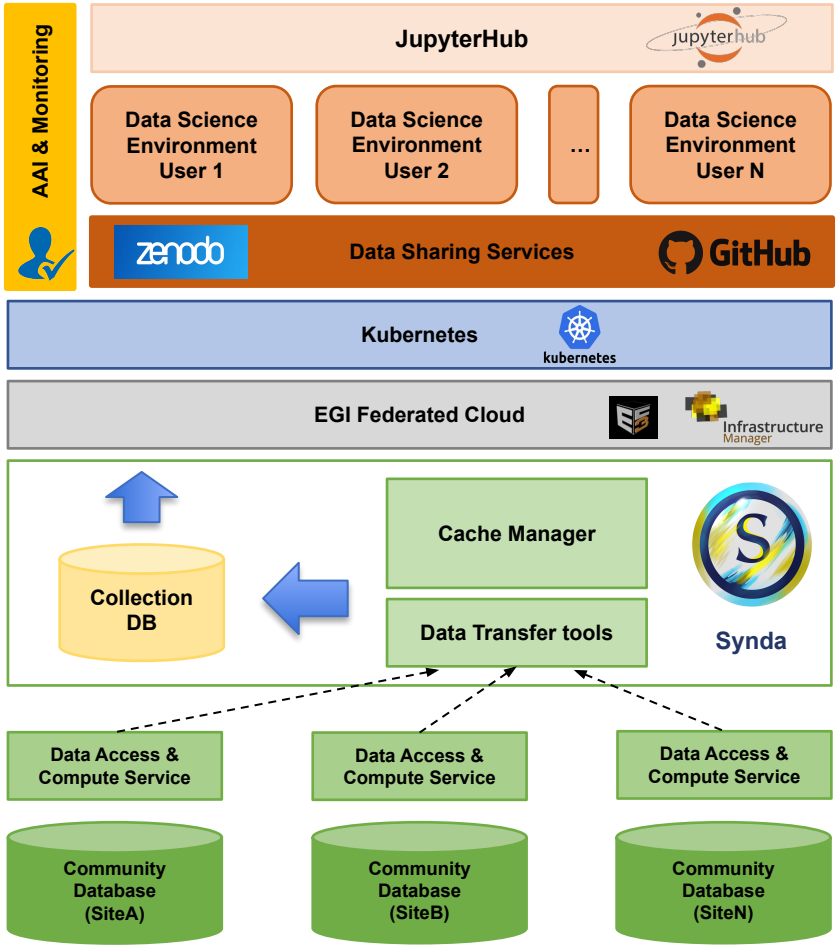
Infrastructure as a Service (IaaS) Cloud

Data collector and Cache Service

Community (legacy) Infrastructure
Community-specific data access services & interfaces

is-enes http://is.enes.org/ ESGF Earth System Grid Federation

The ENES Data Space architecture



EOSC compute services & interfaces

EOSC data sharing services & interfaces

Infrastructure as a Service (IaaS) Cloud

Data collector and Cache Service

Community (legacy) Infrastructure
Community-specific data access services & interfaces

Applications

Application scaling

Data serving

Data

User's feedback from the ENES Data Space

"The ENES Data Space provides access to large data volume and large number of climate simulations for analysis. EGI-ACE was instrumental to calculate simple climate indices on a very large number of simulations."

Christian Pagè,
Research Engineer and Project Manager at CERFACS

"As a system administrator, part of my work is the integration of technologies to solve problems for the users. This is a nice use case to take into account in my workday and the technical part was very interesting to me."

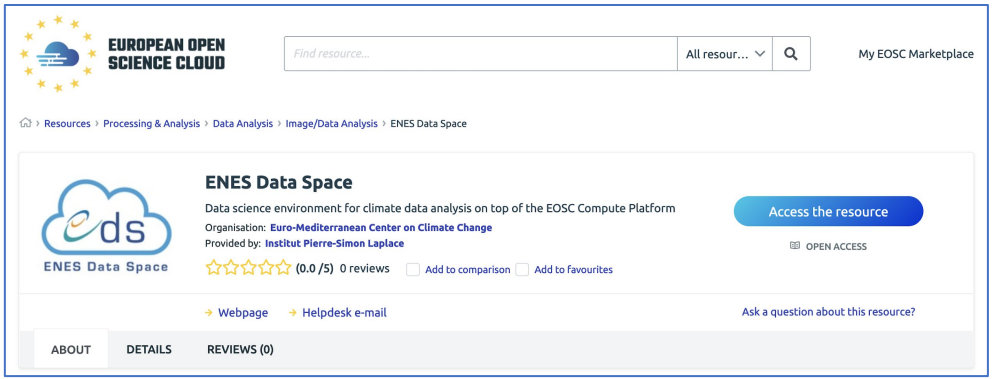
Participant of the ENES Data Space webinar

Status of the ENES Data Space

- Integrated with the **Infrastructure Services** (WP3)
- Integrated with the **Platform Services** (WP4)
- Integrated with the **Federated Services** (WP6)
- Data Space on-boarded in the EOSC Marketplace
- **4** Training events organized + **1** Webinar
- **56** users supported by the Data Space
- **1,633,975** Cloud CPU/h consumed (June 2023)
- Registration of the Data Space in the **IDSA radar as a use case** (in progress)

https://enesdataspace.vm.fedcloud.eu/

enesds-support@cmcc.it



https://marketplace.eosc-portal.eu/services/enes-data-space/

DEMO

Dissemination and Communication

- **25 training events** were organized by the WP5 partners, with focus on domain researchers (M30):
  - **8** by WeNMR,
  - **5** by OpenCoastS,
  - **2** EISCAT_3D,
  - **6** by Virtual Imaging Platform, and
  - **4** by ENES.

- **Additional events organized:**
  - Session at EOSC Symposium 2022
  - Booth at the International Data Space Symposium (2023)
  - Session at European GeoSciences Union (EGU) (2023)

No. of new registered users (M30)

Thematic Service	Baseline (*)	Total users	%
WeNMR	5,625	16,822	+231%
VIP	1,250	1,490	+19%
useGalaxy.eu	25,000	72,669	+92%
OPENCoastS	312	478	+53%
ENES Data Space	0	56	N/A
PROMINENCE	3	8	+113%
LOFAR Science Processing	0	78	N/A
SeaDataNet WebOcean Data Analysis	125	237	+90%
EMSO-ERIC Data Services	250	1,245	+398%
GBIF Cloud data space	10,038	11,828	+18%
Disaster Mitigation & Agriculture (**unfunded**)	187	15	-91%
OPERAS Metrics service and PRISM service (**unfunded**)	N/A	**In development stage**	
OpenRiskNet/NanoCommons Virtual Environment (**unfunded**)	N/A	**In development stage**	

(*) Baseline normalised 15 months before the project

Future plans for the EGI-ACE Thematic Services

EMSO-ERIC Data Services

SeaDataNet WebOcean Data Analysis

useGalaxy.eu

ENES Data Space

EOSC Blue-Cloud2026 — A federated European FAIR and Open Research Ecosystem for oceans, seas, coastal and inland waters

EOSC EuroScienceGateway

EOSC Beyond project (HORIZON-INFRA-2023-EOSC-01-04)

WeNMR Portals suite

VIP

GBIF Cloud Data space

OPENCoastS

LOFAR Science Processing

OPERAS — open scholarly communication in the european research area for social sciences and humanities

OpenRiskNet

iCOMCOT

PROMINENCE

- Fundings from national agencies,
- National infrastructure calls,
- EGI SLA/MoU

KER3 OVERVIEW

ABOUT THE KER

- Thematic Services (Data Spaces and Processing Platforms) from EGI-ACE to the EOSC Exchange
- Leveraging the EOSC Compute Platform (KER1) to deliver scalable environments for data analytics

USER GROUPS

- Researchers
- Scientific communities
- Research Infrastructures

KEY VALUE PROPOSITIONS

- Reduce development time for digital science, and contribute to the cross-fertilisation of knowledge across disciplines
- Facilitating cross-domain sharing and exploitation of research datasets

DISSEMINATION & COMMUNICATION

- **25** training events
- **1,000+** attendees

EXPLOITATION

- **17** Thematic Services from Health Common, Green Deal, Fundamental research and Social Science and Humanities domains
- **~77K** users served
- **77,652,658** CPU/h, **720** TB storage, **333,576** GPU hours delivered

IPR APPROACH

- Open-source licenses Software
- Documentation is licensed under **CC-BY**

Thank you!

Contact: **egi-ace-po@mailman.egi.eu**
Website: **www.egi.eu/projects/egi-ace**

EGI Foundation

@EGI_eInfra

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