Increasing the service offer of the EOSC Portal – InfraEOSC07

EOSC National Tripartite Event Georgia
3rd November 2022

Raul Palma
Poznan Supercomputing and Networking Center (PSNC)
rpalma@man.poznan.pl

RELIANCE project coordinator

The EOSC Future, C-SCALE, DICE, EGI-ACE, OpenAIRE-Nexus and Reliance projects are funded by the European Union Horizon Programme calls INFRAEOSC-03-2020 and INFRAEOSC-07-2020.
InfraEOSC07 overview

- The specific challenge of the topic “Increasing the service offer of the EOSC Portal” (INFRAEOSC07-2020) is to grow the offering available in the EOSC portal with state-of-the-art research enabling services useful to diverse thematic research communities for embracing Open Science practices at the different stages of their research workflows
  - Copernicus - eoSC AnaLytics Engine (C-SCALE)
  - Data infrastructure capacity for the European Open Science Cloud (DICE)
  - EGI Advanced Computing for EOSC (EGI-ACE)
  - OpenAIRE Nexus Scholarly Communication Services for EOSC users (OpenAIRE-Nexus)
  - REsearch LIfeCycle mAnagemeNt for Earth Science Communities and CopErnicus users in EOSC (RELIANCE)
EOSC projects since 2021

coordination, best practices or technology projects (EOSC MAR 2023-24)
Increasing the EOSC exchange service offer

365 Resources onboarded in EOSC as of end of October 2022 (source EOSC Metrics Portal)

EOSC Horizontal services are delivered (e.g. data, compute and other research enabling services)

The onboarding process include validation of data sources

Ability to create thematic execution environments / VREs based on integration of compliant thematic, horizontal, and core resources

Horizontal services and resources from 07 projects and e-infras onboarded

Cluster services onboarded

Thematic clusters

Regional clusters

INFRAEOSC-07
EOSC provisioning / expansion
The C-SCALE Project

Europe lacks an integrated compute and storage infrastructure for the exploitation of Copernicus datasets in scientific and applied applications.

C-SCALE responds to that challenge by enhancing the EOSC Portal with pan-European federated data and computing infrastructure services for Copernicus.

C-SCALE: Copernicus - eoSC AnaLytics Engine
- Project duration: Jan 2021 – June 2023 (30 months)
- Budget: ~ 2 million Euros
- Consortium of 11 partners with pan-European coverage
Key project results

C-SCALE services enabling European researchers to easily discover, access, process, analyse and share Copernicus data, tools, resources and services through EOSC – they will be onboarded on EOSC Marketplace before the end of the project:

FedEarthData
Uniform access to a federation of computing and data providers to execute Copernicus and Earth Observation workloads.

openEO Platform
Versatile cloud-based processing and analytics environment for Earth Observation data on infrastructures supporting openEO API

Metadata Query Service
One-step discovery of Copernicus data in data archives across the C-SCALE Data Federation and across Europe

C-SCALE Community
Functional co-design with scientific communities across Europe through a user forum for Copernicus users accompanied by a documentation site.

C-SCALE Federation\(^1\) is driven by user requirements to guarantee the delivery of an environment that satisfies user needs:

- 6 project-internal use cases deployed on the C-SCALE federation
- **C-SCALE Open Call**: 14 external use cases (6 are already deployed and 8 are in the progress of onboarding)

\(^1\)An open compute and data federation as an alternative to monolithic infrastructures for big Earth data analytics

Workflow solutions
Easy deployment of workflows supporting monitoring, modelling and forecasting of the Earth system
Project summary

Consortium: 24 Partners  
Coordinator: CINECA

30 months  
(1 Jan 2021 – 30 Jun 2023)

Grant Agreement: 101017207  
Call: INFRAEOSC-07-2020

Enable a European storage and data management infrastructure for EOSC  
Provide generic services to store, find, access, and process data in a consistent and persistent way

Budget: € 6.997.706,66

18 providers from 11 European countries are offering 14 state-of-the-art data management services together with more than 50 PB of storage capacity
Project outcomes

14 offered services covering the whole research data management lifecycle

- Services are horizontal and can be used in any scientific domain
- Back-end resources are offered free at the point of use for the project duration
- Support is provided to users for customised solutions
- Services are available via the EOSC marketplace (here)

Recent developments include (among others)

- Integration of operation services with EOSC (AAI, helpdesk, onboarding)
- Long term preservation policies definition
- Improvements in connecting data services with HPC resources
- PID integrity check
- Sensitive data management
- Use case demonstrators (including integration with the other 07s’ services)

For more info see the project website (https://dice-eosc.eu/deliverables)

Services categories

- **Data Discovery**
  - Making research data findable
  - Data source harvestable and to make data widely discoverable

- **Data Repository**
  - Long-term data resources preservation and sharing
  - To store non-active FAIR data

- **Policies based Data Archive**
  - Mid-long term data resources, accessible from computing facilities to store long-term non-active research data
  - Value-added services (integrity checks, replications, ..)

- **Data Archive**
  - Mid term storage resources after/between projects, Accessible from computing facilities to store non-active research data, bitwise preservation

- **Personal and/or Project workspace**
  - Mid-term storage during project,
  - Good connection to computing facilities,
  - To store active resource data
Consortium:
- Coordinator – Stichting EGI
- 33 Partners, 23 third parties
- Most of EGI council members + Several Research Infrastructures (~ERICs)

Scope:
- 49% service delivery (Virtual Access)
- Co-development of ‘thematic services’ with research communities

Duration:
- Jan 2021 – June 2023 (30 months)

https://www.egi.eu/project/egi-ace/
Main activities and achievements

1. Service delivery in EOSC:
   - 35 services from the consortium
   - 8 services from external providers – empowered by our services

2. Integration of Thematic Services:
   - 15 Thematic Services (VREs, Data Spaces) served
   - 78,100 users

3. Onboarding and serving user communities:
   - Over 150 user communities
   - 35M+ cloud CPU/h consumed
   - 17 webinars, 500+ attendees, 900+ YouTube views
OPENAIRE
A European infrastructure on open scholarly communication since 2009

Mission: Shift scholarly communication towards openness and transparency and facilitate innovative ways to communicate and monitor research.

Non-profit organization
- Established Oct 2018
- Headquarter Greece, virtual office
- 48 members
- From 34 countries
19 Scholarly communication services for all

Covering all research lifecycle

Manage Data
Publish
Discover
Interoperability
Outreach
Assess

http://catalogue.opinaire.eu
Some numbers (this year to date)

- **AAI**: 24K users actively using OpenAIRE AAI (*most services do not require login*)
- **Zenodo**: 181K researchers have shared 66PB of data
- **Amnesia**: 15K visits, 450 downloads (installations?)
- **Argos**: 1200 researchers making their DMPs
- **PROVIDE**: 2.5K data sources registered & validated, generating 25M brokering events
- **UsageCounts**: 500 repositories capturing 785M downloads
- **EXPLORE**: 1700 users claimed 61K research results to 1700 projects and 35K claims to ORCID
- **OpenCitations**: 1.3B citation records, 82M API requests
- **MONITOR & CONNECT**: 40+ dashboards in use
- **Episciences**: 25 Journals, 10K users and 500 submissions
- **Graph**
  - 145M publications, 18M research data, 304K research software items, from 110K data sources, linked to 3M grants and 178K organizations
  - 35M API requests, 6K dumps
Reliance at a Glance

• RELIANCE is extending the EOSC's capabilities with an enhanced support for research activities via a set of interconnected services, in alignment with the EOSC IF.

• It aims to enhance the discovery of and access to research data/results (incl. Copernicus), improve the extraction of relevant information, manage the research lifecycle while promoting FAIR and open science principles.

• It is demonstrating its services' value via 3 Earth Science communities and others via an Open Call, fostering the use of Copernicus data, enhancing EOSC support for multidisciplinary research and improving EU science as a whole

• Project duration: Jan 2021 – June 2023 (30 months)
• Budget: ~ 2 million Euros
• Consortium of 9 partners with pan-European coverage
Core services
Research Lifecycle Management Ecosystem Pillars

- **Research Objects** as the overarching mechanism to manage scientific research activities and connect associated resources

![RoHub](http://www.researchobject.org)

- **Data Cubes** for efficient and scalable structured data access and discovery

![ADAM](http://www.researchobject.org)

- **Text mining** services to extract machine-readable metadata enabling researchers to discover scientific information at scale and to structure their own research.

![EOSC](http://www.researchobject.org)

http://www.researchobject.org
Key project outcomes

- RELIANCE services are onboarded in EOSC (5 production services). 7 other added-value and/or experimental services are also available and/or being integrated.
- ~100 users
- ~3K Research Objects, aggregating
  - 12K+ resources
  - 48K+ annotations
- 3 user communities, with 5 grand challenges and use cases, including a multidisciplinary one, and 5 other engaged via the RELIANCE open challenge
  - Sea monitoring
  - Climate change
  - Geohazards
- 2 other use cases from EOSC DIH call for pilots
- RELIANCE services interconnect and integrate several other EOSC services (8 so far), making it one of the best examples enabling researchers to fully exploit EOSC:
InfraEOSCo7 collaboration

- 07 projects have a strong collaboration, which was made official with a multi-lateral agreement signed by the five projects and EOSC-Future, covering a set of topics:
  - technical interoperability and integration,
  - joint outreach and dissemination activities,
  - training activities
- The projects coordinate and/or support different EOSC WG like the Semantic Interoperability, Research Products Publishing FW, Compute continuum, etc.
- The projects also collaborate to support the implementation of the EOSC Future Science Projects of the thematic clusters ESCAPE, PaNOSC, SHOCC, EOSC-Life, ENVRI-FAIR.
- The projects support each other and collaborate in the definition and implementation of the Virtual Access assessment for EOSC
- The projects support together many other activities like AMA sessions, MAR, horizontal services identification, etc.
- The projects participate and support the EOSC DIH pilot activities
InfraEOSCo7 collaboration on technical activities

1. Architecture & Technical Interoperability
   - Evolve the EOSC Architecture
   - Contributions to the EOSC Core design and development
   - Contribution to the interoperability guidelines for EOSC Core
   - Contribution to the interoperability guidelines for EOSC Exchange
   - Make services and resources compliant to the EOSC Interoperability Framework guidelines
   - Resource Composability Demonstrators and Early Adopter pilots

2. Service Provisioning
   - Resource Onboarding
   - Selection of interdisciplinary services to create a set of EOSC horizontal services

3. Resource Provisioning and Technical Support
   - EOSC Portal as aggregation channel
   - Identification of and access to the EC-funded resources
   - Resources for Science Projects, DIH business pilots and other use cases and technical support
InfraEOSC07 collaboration on uptake activities

1. Joint promotional activities
   a. joint articles to highlight how the projects are complementary and how differently contribute to the EOSC Ecosystem;
   b. coordinated promotion of the projects’ open calls
   c. development of an EOSC provider toolkit that includes ready-made materials explaining why and how to become an EOSC provider.
   d. Promotion of EOSC uptake success stories
   e. Promotion of training and training materials developed by the EOSC Future and the INFRAEOSC-07-2020 projects
   f. Coordinated population and update of the EOSC Portal website

2. Joint engagement activities through events
   a. “Ask me anything” sessions
   b. EOSC Future provider days
   c. EOSC hackathon
InfraEOSCo7 collaboration on training activities

1. Joint EOSC training activities
   a. A series of joint training activities (at least 2 per year)
   b. EOSC Training Roadmap as a living document
   c. A collection of reusable training materials available for EOSC actors

2. Identify existing and co-development of new EOSC training materials, and building the EOSC Future Knowledge Hub
   a. Create an overview of training materials available for re-use
   b. Combine existing training materials in the training packs and repackage them if needed
   c. Co-develop new training materials to fill in the gaps
   d. Feature/host (where relevant) training materials in the EOSC Future Knowledge Hub
Example: collaborating research enabling services in support of open science

Find research work, access and reproduce it, reuse it in new research, collaborate, assess quality and publish it leveraging different EOSC services.
Example: collaborating research enabling services in support of open science

**Research Object:** account, describe and share everything about your research, including how those things are related

http://www.researchobject.org

The RO has its own metadata, can be managed and evolved in its own right, and it can be packaged, deposited, transferred, accessed, and reproduced if appropriate
Example scenario (EOSC)

Find research work, access and reproduce it, reuse it in new research, collaborate, assess quality and publish it leveraging different EOSC services.
See you next time!

Raul Palma
rpalma@man.poznan.pl

The EOSC Future, C-SCALE, DICE, EGI-ACE, OpenAIRE-Nexus and Reliance projects are funded by the European Union Horizon Programme calls INFRAEOSC-03-2020 and INFRAEOSC-07-2020.