Reading list: User & Resource Environments

EOSC Association

The EOSC Association is the legal entity established to govern the European Open Science Cloud (EOSC). It was formed on 29th July 2020 with four founding members and has since grown to more than 250 Members and Observers. The Association membership is jointly responsible for delivering the objectives agreed in the Memorandum of Understanding signed by the European Union and EOSC Association to form the official Partnership.

Webpage: https://eosc.eu

Strategic Research and Innovation Agenda(SRIA) of the European Open Science Cloud (EOSC) - Version 1.2

The overall purpose of the EOSC Strategic Research and Innovation Agenda, or SRIA, is to define the general framework for future research, development and innovation activities in relation to the European Open Science Cloud. The current version of the SRIA (1.2), including MAR 2025-2027, was approved by the EOSC Partnership Board at their 5th meeting on 06 December 2023.

Link: https://eosc.eu/wp-content/uploads/2023/12/20231114_SRIA_1.2_final2.pdf

EOSC Task Forces

The EOSC Task Forces address key areas in the implementation of the European Open Science Cloud (EOSC), providing feedback on developments, identifying strategic gaps, and suggesting areas for investment.

Webpage: https://eosc.eu/eosc-task-forces/

List of EOSC-related Projects (Horizon Europe Projects)

This list includes European projects that support the implementation and development of the EOSC.

Link: https://eosc.eu/horizon-europe-projects/

Task Force: Technical Interoperability of Data and Services

The Technical Interoperability of Data and Services Task Force will take the EOSC Interoperability Framework (EIF) recommendations around technical architecture as their starting point to help develop the EOSC Core and Exchange as described in the SRIA. The TF will deliver key documents such as a first principles document and a landscape overview of the EIF, as well as technical architecture descriptions of the EIF, including examples of adaptation hints for major existing solutions.

Webpage: https://eosc.eu/advisory-groups/technical-interoperability-data-and-services

Design Considerations for Technical Interoperability in EOSC

The document introduces design-stage principles, both architectural and non-architectural, that are commonly met in software design and maintenance best practices, but are in this case approached from the perspective of EOSC's technical interoperability. These 21 principles are additionally catalogued in Annex A in tabular form including the guidelines and recommendations of this report, together with the suggested criteria for satisfying the respective principles.

The document then proceeds to specify suggestions for the interoperability of EOSC services, addressing the pillars of Open Science, Cloud Computing and the User. The role of standards, as

Funded by the European Union

the cornerstone of system interoperability, is presented briefly, and Annex B contains an evidence-based enumeration of the relevant specifications. In the Annex, an indicative list of standard families, along with sources where common standards are drawn for EOSC services, is presented. In conclusion, a set of key points is summarised, that emphasises the need to adopt an interoperability strategy at the time of design of a given system.

Link: https://zenodo.org/records/8109528

A landscape overview of the EOSC Interoperability Framework - Capabilities and Gaps

The document summarises all the work that has been done on the EOSC Architecture and Interoperability Framework by different initiatives in the last years that allowed the convergence towards a common idea of what the EOSC Interoperability Framework should be. In line with the concept of System of Systems, the EOSC Interoperability Framework aims to put together, in a unique framework, the results and the decades of experience on interoperability of ESFRI/Thematic clusters, ERICs, general purpose e-Infrastructures (EGI, EUDAT, GEANT, OpenAIRE) and cross-domain initiatives like the Research Data Alliance, GO-FAIR and CODATA. The document presents the foundations of the EOSC Interoperability Framework, the EOSC Architecture and the Minimum Viable EOSC as they have been defined by the EOSC Working Groups, established during the previous EOSC Governance, and describes their evolution under EOSC Future, other relevant EOSC projects and via the different task forces of the EOSC Association.

Link: https://zenodo.org/records/8399710

EuroScienceGateway (project)

EuroScienceGateway will leverage a distributed computing network across 13 European countries, accessible via 6 national, user-friendly web portals, facilitating access to compute and storage infrastructures across Europe as well as to data, tools, workflows and services that can be customised to suit researchers' needs. At the heart of the proposal workflows will integrate with the EOSC-Core. Adoption, development and implementation of technologies to interoperate across services, will allow researchers to produce high-quality FAIR data, available to all in EOSC. Communities across disciplines — Life Sciences, Climate and Biodiversity, Astrophysics, Materials science — will demonstrate the bridge from EOSC's technical services to scientific analysis.

Webpage: https://galaxyproject.org/projects/esg/

Galaxy Community Hub

Galaxy is an open-source platform for FAIR data analysis that enables users to:

- Use tools from various domains (that can be plugged into workflows) through its graphical web interface.
- Run code in interactive environments (RStudio, Jupyter...) along with other tools or workflows.
- Manage data by sharing and publishing results, workflows, and visualisations.
- Ensure reproducibility by capturing the necessary information to repeat and understand data analyses.

The Galaxy Community is actively involved in helping the ecosystem improve and sharing scientific discoveries.

Link:

- https://galaxyproject.org
- https://rdmkit.elixir-europe.org/galaxy_assembly

Funded by the European Union

Galaxy: a System to compose VRE

- Galaxy can be considered a VRE with support of the full data life-cycle.
  Link: https://rdmkit.elixir-europe.org/galaxy_assembly
- Galaxy is offering a plug-in system to extend its functionality (e.g. for data importers, data exporter, object stores, tools, visualisations, AAI, user-management ….).
- List of all tools offered at the European Galaxy server
  Link: https://usegalaxy-eu.github.io/tools
- Galaxy can manage Jupyter Notebooks, RStudio instances, Shiny applications and all other systems that are web-accessible - we call them **Interactive tools (ITs)**
- **Interactive tools** can be integrated into Galaxy by creating 1) a container (Docker, Podman etc.) and 2) a metadata description of the service. Metadata contains technical ones, like exposed port of the container, but also input (and datatypes) and outputs for provenance capturing, parameters to change runtime behaviour of the application etc.
- **Interactive tools** can be scheduled to arbitrary clusters and job scheduling systems, including SLURM, htcondor, k8s and cross cloud. For example, a Galaxy in Italy can schedule **Interactive tools** in Spain and everywhere else.
- Data can be stored in Galaxy (with all its sharing and data management capabilities) and injected into **Interactive tools.** Data created inside the **Interactive tool** (e.g. images from a Jupyter Notebook, or the Notebook itself) can be transferred back to Galaxy - this combines classical HPC workflows with Jupyter like data processing
  Link: https://galaxyproject.org/news/2020-04-14-integrative-meta-omics/
- The source of a few **Interactive tools** can be found on github and deployed at https://live.usegalaxy.eu and other Galaxy servers.
  Link github:
  https://github.com/usegalaxy-eu/galaxy/tree/release_23.0_europe/tools/interactive
  Link Galaxy server: https://live.usegalaxy.eu/
- We offer trainings to integrate ITs/VREs/tools into Galaxy:
  - Galaxy Tool integration
    Link: https://planemo.readthedocs.io/en/latest/writing.html
  - Specific IT training
    Link:
    https://training.galaxyproject.org/training-material/topics/dev/tutorials/interactive-tools/tutorial.html

EGI (organisation)

We are the federation of computing and storage resource providers united by a mission of delivering advanced computing and data analytics services for research and innovation. Our research services include high-throughput and cloud computing, storage and data management, identity and access management, analytics, consultancy and support, training and co-development. We also provide a wide range of internal services for our federation and for our partners from the business world.

Webpage: https://www.egi.eu

EGI Service Catalogue

Catalogue of services offered by EGI.

Funded by the European Union

Link: https://cdn.egi.eu/app/uploads/2023/06/egi-service-catalogue-digital-spreads.pdf

Background reading material

A list of materials not produced by the programme committee parties but of importance to the EOSC Winter School track User & Resource Environments.

Project AI4EOSC: demonstrator

AI4EOSC project will deliver an enhanced set of services for the development of AI, ML and DL models and applications in the EOSC. The services will make use of advanced features such as distributed, federated and split learning; provenance metadata; event-driven data processing services or provisioning of AI/ML/DL services based on serverless computing.
Link: https://ai4eosc.eu
AI4EOSC virtual environment: https://dashboard.cloud.ai4eosc.eu/marketplace

Project FAIR-EASE: demonstrator

FAIR-EASE focuses on opening gateways for the earth and environmental sciences which require a large volume of data from sources such as satellite, in-situ observations, models, and even omics experiments. Earth system domains are interconnected, but the present digital architecture is based essentially on distributed and domain-dependent data repositories inducing real difficulties for integrated uses of all the environmental data. This creates barriers for several studies with large societal impacts, such as climate change, agriculture and food, human safety and health. To go beyond this state-of-the-art, the overall objective of FAIR-EASE is to customize and operate distributed and integrated services for observation and modelling of the Earth system, environment and biodiversity by improving their different components implemented in close cooperation with user-communities, the European Open Science Cloud (EOSC) and research infrastructures in their design and sustainable availability.
Link: https://fairease.eu/about
FAIR-EASE virtual environment: https://earth-system.usegalaxy.eu/
FAIR-EASE virtual environment: https://ecology.usegalaxy.eu/

Project RAISE: in development phase of new VRE

RAISE project aims to provide the mechanisms for a distributed crowdsourced data processing system, moving from open data to data open for processing. To do so, RAISE will attempt to adapt open data to the culture of the research community, ensuring FAIR principles. The vision of the project is the EOSC Web of FAIR Data and Services for Science is an open, fair and reliable Research Community where every researcher will be accredited for their work and all research data will be equally accessible for processing without violating data protection regulations.
Webpage: https://raise-science.eu/publications/
Virtual environment planned for April 2024

Funded by the European Union

Project AquaINFRA: in development phase of new VRE

The AquaINFRA project aims to develop a virtual environment equipped with FAIR multi-disciplinary data and services to support marine and freshwater scientists and stakeholders restoring healthy oceans, seas, coastal and inland waters. The AquaINFRA virtual environment will enable the target stakeholders to store, share, access, analyse and process research data and other research digital objects from their own discipline, across research infrastructures, disciplines and national borders leveraging on EOSC and the other existing operational dataspaces. Besides supporting the ongoing development of the EOSC as an overarching research infrastructure, AquaINFRA is addressing the specific need for enabling researchers from the marine and freshwater communities to work and collaborate across those two domains.

Webpage: https://aquainfra.eu

Deliverables: https://aquainfra.eu/deliverables

- D2.1 AquaINFRA Architecture
- D3.1 AquaINFRA DDAS user requirements

AquaINFRA virtual environment: https://aqua.usegalaxy.eu/

Funded by
the European Union