EOSC-Synergy: Integrating Capacities & Building Capabilities

Jorge Gomes (LIP/INCD) on behalf of the EOSC-Synergy consortium

Horizon 2020 Research and Innovation Programme
Grant Agreement # 857647

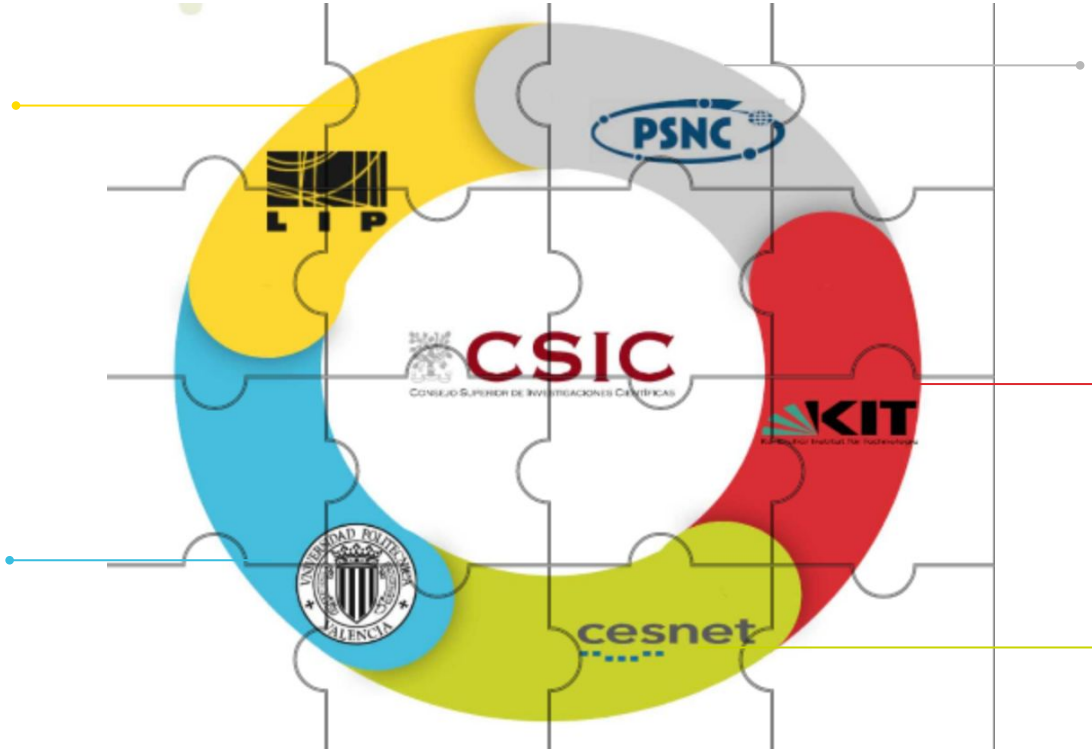
Implementing EOSC at national level

Promoting EOSC High Quality Services

Software quality as a service, FAIRness evaluation and quality certification badges.

Thematic Services Integration

10 thematic services addressing 4 scientific areas (Earth Observation, Environment, Biomedicine and Astrophysics).



Skills development

Environment for tutorials with a dedicated MOOC platform, courses methodology and a Hackaton as a service platform.

Capacity Expansion at the Infrastructure level

Integration of services and resources from the RIs of the consortium partners.

Alignment at the Policy Level

Collaboration with regional projects on landscaping activities, gap analysis and contribution to EOSC policies.

Project partners

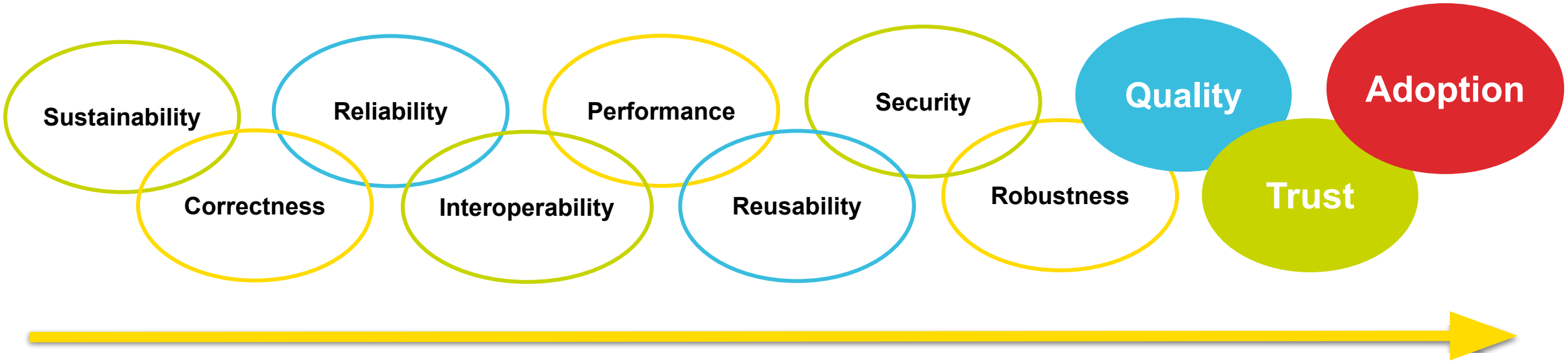
EOSC-synergy coordination structure is based on IBERGRID: www.ibergrid.eu

Project management and legal coordination resides on CSIC (Spanish coordinator of IBERGRID)

Spain, Portugal, UK, Czech Republic, Germany, Slovakia, Poland and the Netherlands

Fostering Service Integration and Adoption

Quality based approach for service integration to promote EOSC adoption

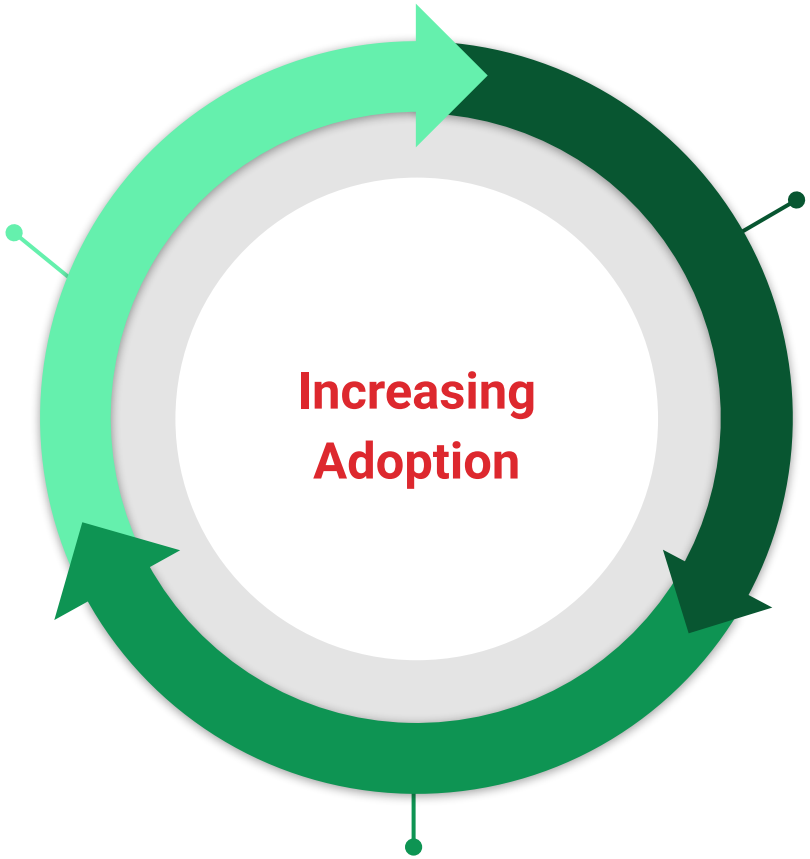


Virtuous cycle

Improved Software Quality

1-Developers/Integrators

Increase software quality
Adhere to software development best practices



2-Providers/Operators

Increase service quality
Adhere to service delivery best practices

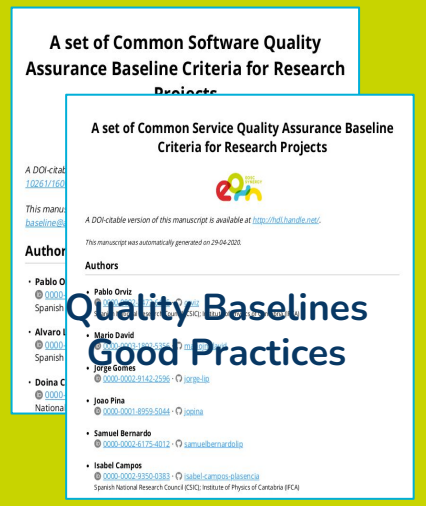
Increased Usage

Improved Services Quality

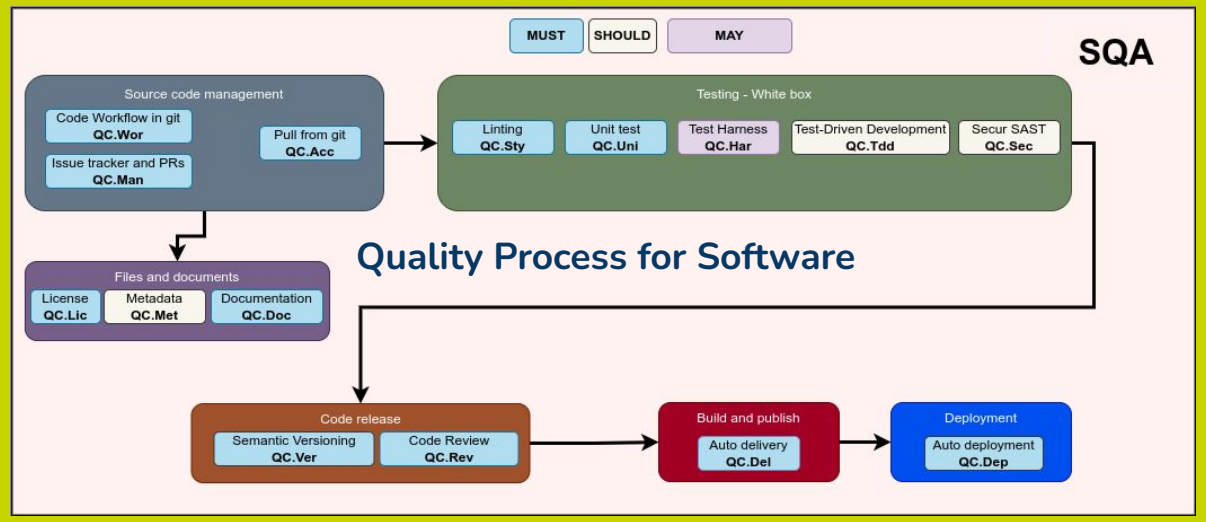
3-Users/Researchers

More aware of EOSC services quality
Build trust and increase adoption

SQAaaS Platform



Quality Baselines
Good Practices



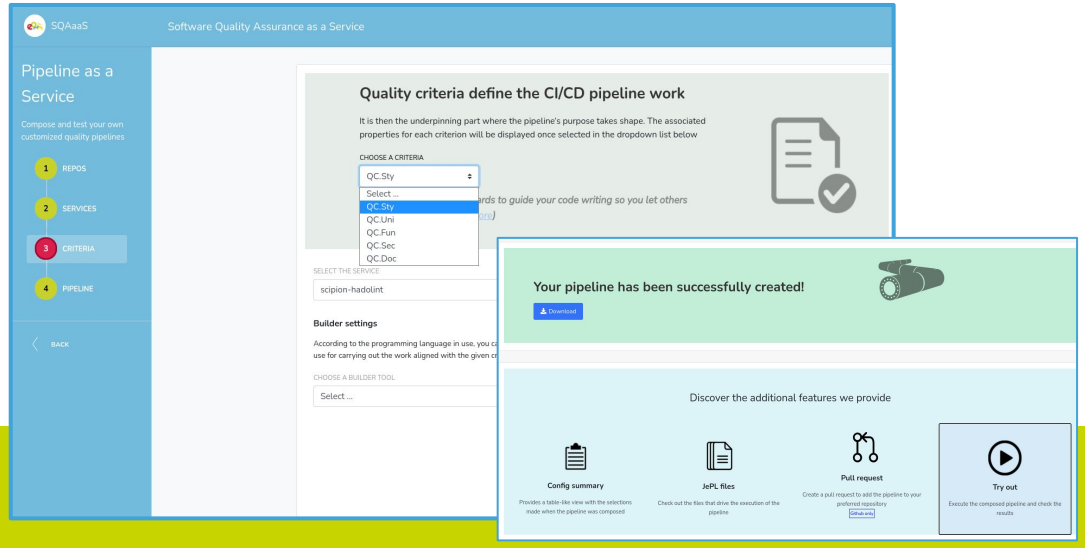
Software Classification			
Quality Criteria	Quality Badges		
	Gold	Silver	Bronze
QC.Acc	■	■	■
QC.Lic	■	■	■
QC.Sty	■	■	
QC.Met	■	■	■
QC.Uni	■	■	
QC.Doc	■		■
QC.Sec	■	■	
QC.Wor	■		
QC.Ver	■	■	
QC.Man	■		
QC.Del	■		

Mapping of QC to Badges

SQAaaS Platform leverages CI/CD & DevOps

Source Code Repository

GitHub

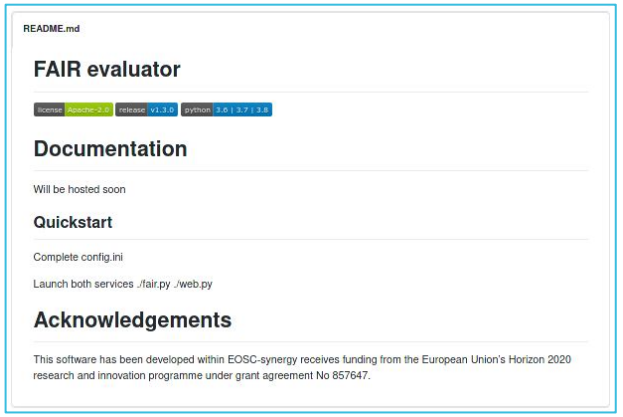


Badges to Reward Quality

Tools for FAIR assessment

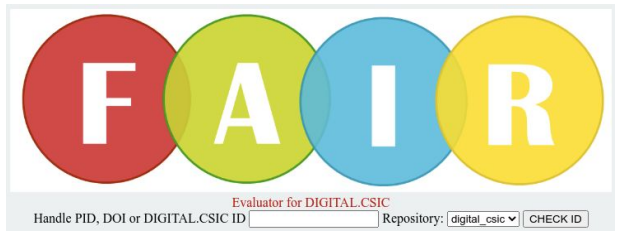
FAIR EVA

- Assess FAIR compliance of research data
- Implements the RDA indicators
- Modular architecture supporting multiple types of data repositories



Integration with the SQAaaS

- FAIR validation in Quality Assurance
- Supporting assessment tools in CI/CD pipelines (FAIR EVA and F-UJI)
- Support for FAIR badging



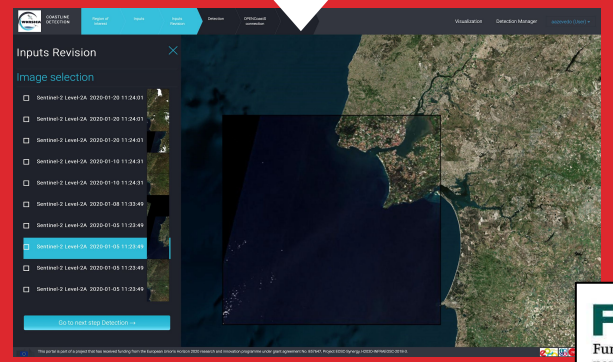
https://github.com/EOSC-synergy/FAIR_eva

Thematic Services in Earth Observation

WORSICA

Water Monitoring Sentinel Cloud Platform

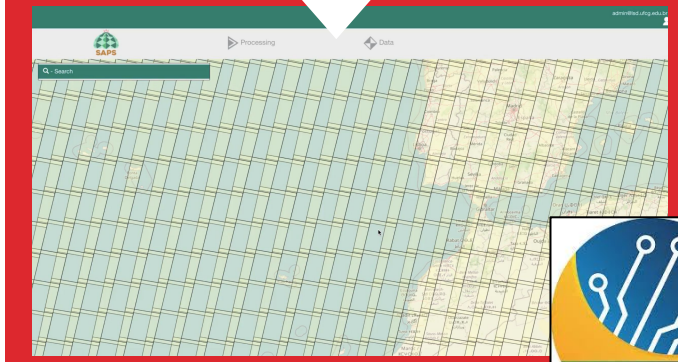
A service for the detection of water using satellites, Unmanned Aerial Vehicles & in-situ data. WORSICA can be used for coastline detection, inland water bodies detection and water leaks detection on irrigation networks.



SAPS

Surface Energy Balance Automated Processing Service

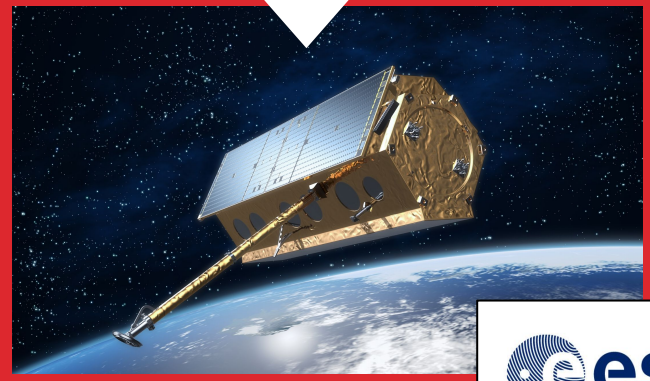
Used to estimate Evapotranspiration and other environmental data that can be applied, for example, on water management and the analysis of the evolution of forest masses and crops.



GCore

Acquisition, cataloguing and processing EOS data

G-Core is a production-ready technology used as a service at ESA's and national programs that provides a Data Manager for spatial and non-spatial purposes and a framework for third-party processors.



Thematic Services in Biomedicine & Astrophysics

SCIPION

CryoEM data processing for Structural Biology

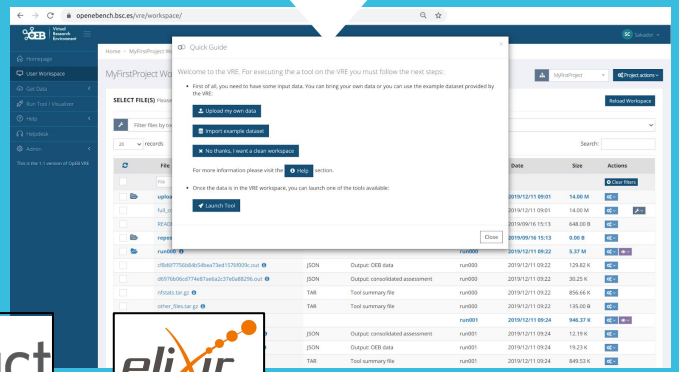
ScipionCloud service will allow users from Instruct to deploy a dynamic cluster in the cloud to keep processing the data acquired at the facility.

EIRENE

OpenEBench

ELIXIR benchmarking and technical monitoring platform

Used to evaluate bioinformatics tools, OpenEBench is an observatory for SW quality based on the automated monitoring of FAIR for research software metrics and indicators.



LAGO

Latin American Giant cosmic ray Observatory

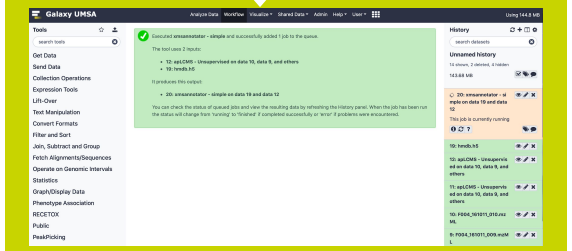
LAGO is a cosmic ray observatory made of a network of water- Cherenkov detectors (WCD) spanning over different altitudes and latitudes making research on High Energy Physics, Space weather, etc.

Thematic Services in Environment

UMSA

Untargeted Mass-Spectrometry Analysis

UMSA aims at processing data to correlating the whole spectra with other data to work with more complex hypotheses on the impact of environment in human health.



EIRENE

MSWSS

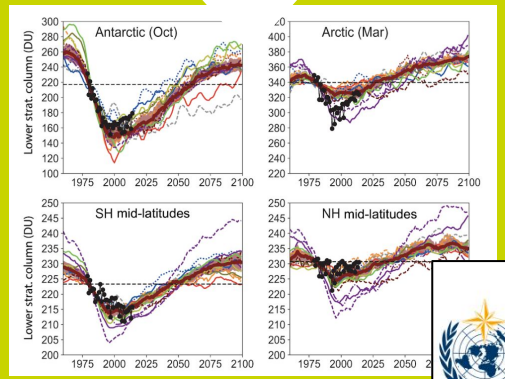
Water Supply Systems modeling and analysis

MSWSS integrates the analysis and simulation of toxics in drinking water supply networks to allow operators and researchers to analyse hazardous events.

O3AS

Ozone Analysis Service

The O3AS service shall provide an invaluable tool to extract O_3 trends from large climate prediction model data to produce figures of stratospheric ozone trends.

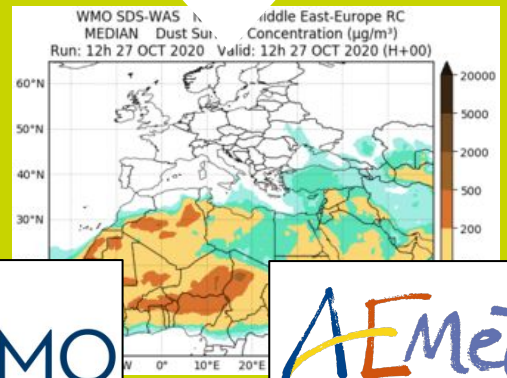


WMO

SDS-WAS

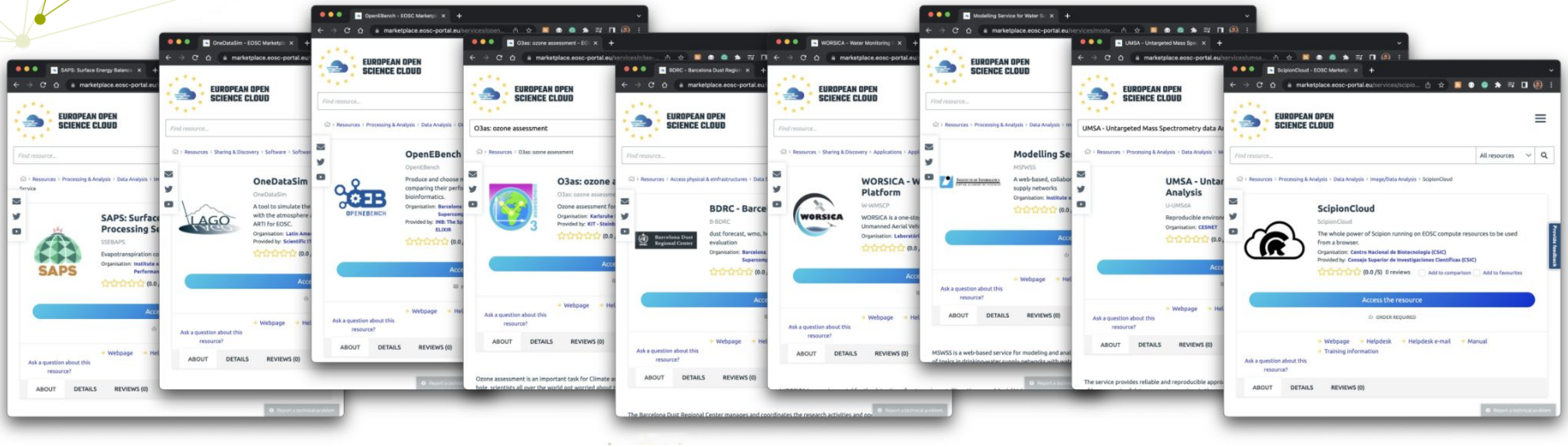
A Service related to the mineral dust forecast

SDS-WAS aims to support institutional entities to warn about possible dust events and to foster the study of dust-related phenomena.



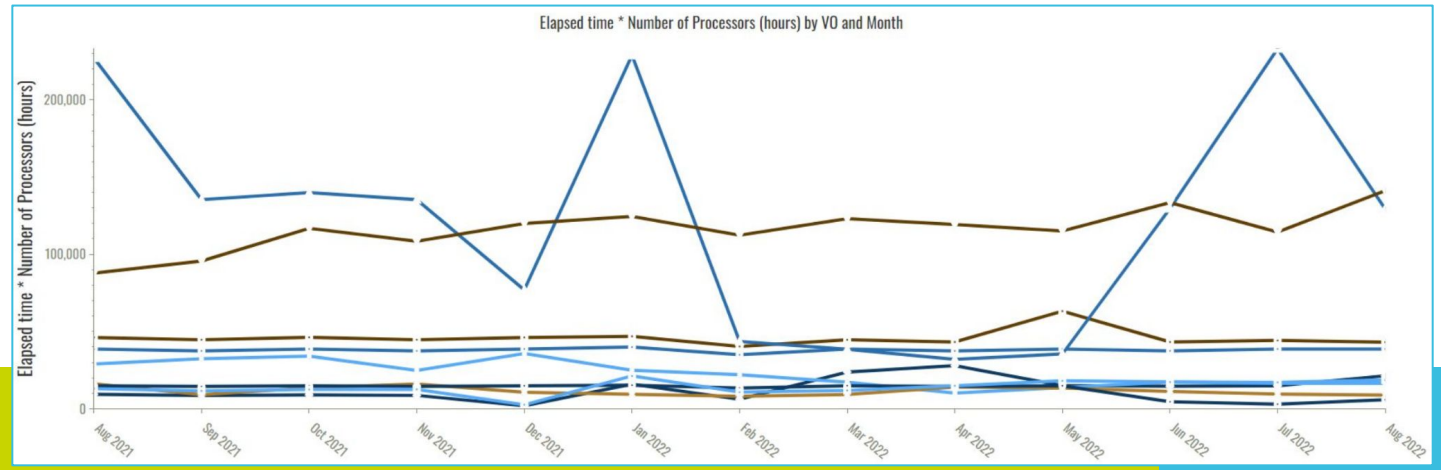
Agencia Estatal de Meteorología

Increasing users and usage



Over last year:

- More than 4.7 Million CPU core hours
- More than 3.400 VMs
- 90 registered users in VO



Infrastructure for thematic services

	WORSICA	G-Core	SAPS	Scipion	LAGO	SDS-WAS	UMSA	MSWSS	O3AS	OpenE Bench
AAI	EGI Check in	Kerberos LDAP & CAS User/pwd	EGI Check in	EGI Check in	eduTEAMS+ EGI Check in	B2ACCESS	EGI Check in +Life Science AAI	EGI Check in	EGI Check in	Life Science AAI
Workload Manager	ArcCE, Batch (SLURM)	GCore+ K8s	K8s	Batch (SLURM)	Batch (SLURM)	Batch (SLURM)	Batch (SLURM) in IM/EC3 (in Galaxy)	Batch (SLURM) in EC3 (in Galaxy)	Cluster batch (SLURM) & K8s	GA4GH WES/TES stack + NextFlow
Resource Manager	IM (TOSCA)	IM / EC3	IM / EC3	IM / EC3	Local clusters & IM+EC3	EC3	IM / EC3	IM / EC3	IM	one
Data Storage	Nextcloud, Dataverse	ElasticSe-arch for the catalogue	Open Stack Swift	Local + S3	EGI DataHub ONEDATA	B2HANDLE /B2SAFE	Local + S3	Local + ONEDATA	WebDAV	Local + B2SHARE

Integrate computing and storage infrastructure

EOSC-SYNERGY Handbook

https://handbook.eosc-synergy.eu/

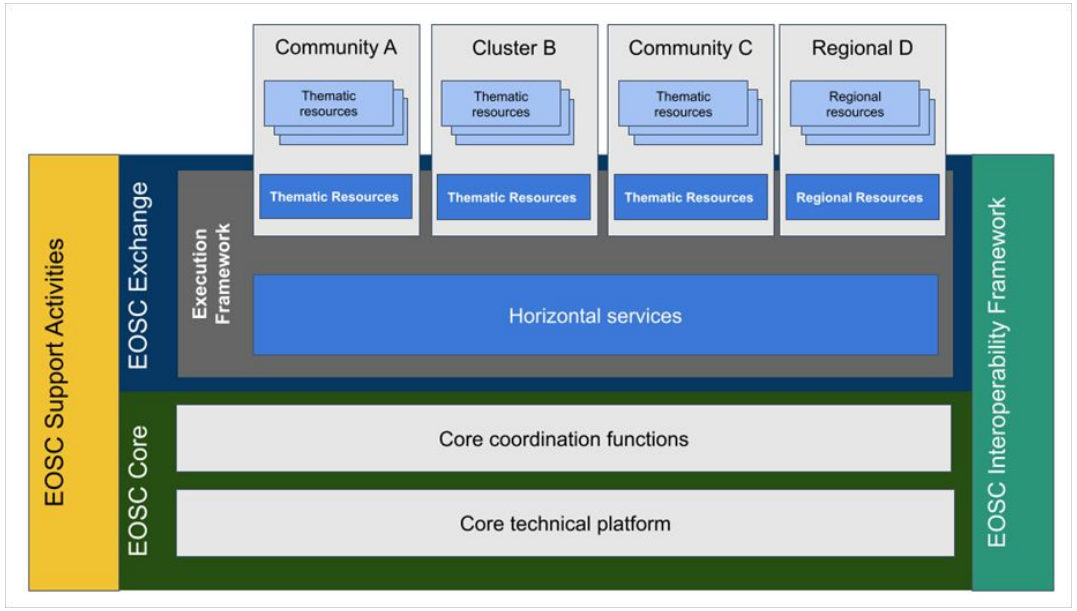
EOSC SYNERGY

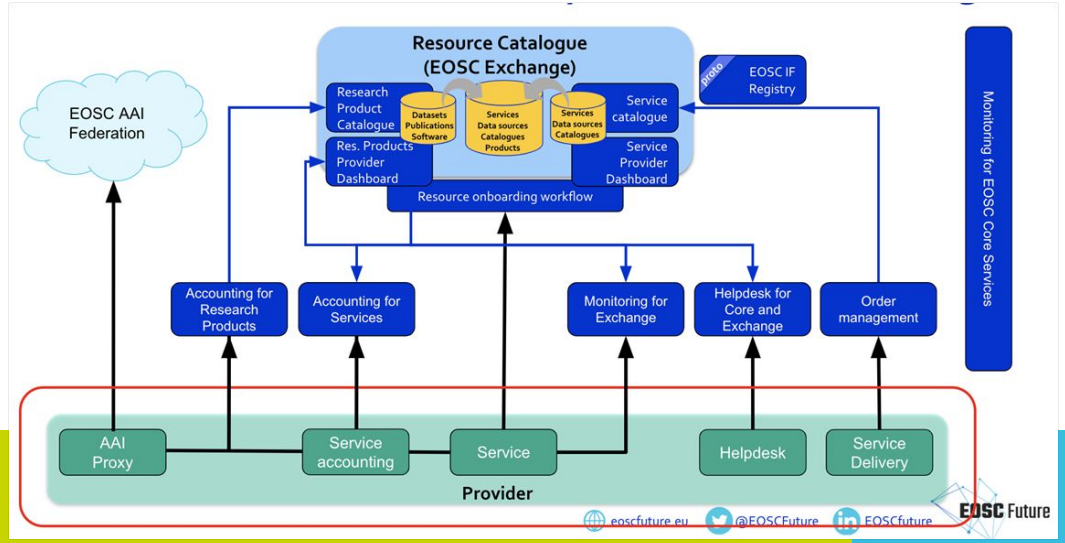
EOSC-SYNERGY – Handbook

This handbook is targeted at multiple different audiences. Different sections are targeted at different groups. We suggest a prioritised list to read for the different target audiences.

- **Management of Computer Centres**:
  a. Introduction
  b. Architecture of EOSC
  c. Resource Integration
- **Users of the cloud**:
  a. Introduction
  b. Thematic Services
  c. EOSC Synergy Services and tools
- **System Administrators**:
  a. Resource Integration
  b. Thematic Services
  c. Architecture of EOSC

Imprint





Policy related activities

Recommendations for the alignment of national policies related to EOSC.

Aimed at policy makers, EOSC Bodies, research funders, research performing organisations

https://www.eosc-synergy.eu/policy-harmonization/

Expanding training and education capabilities through an innovative online platform

Guidance for creating good quality tutorials

- Best practices and training related materials

Set of EOSC ecosystem tutorials and training materials

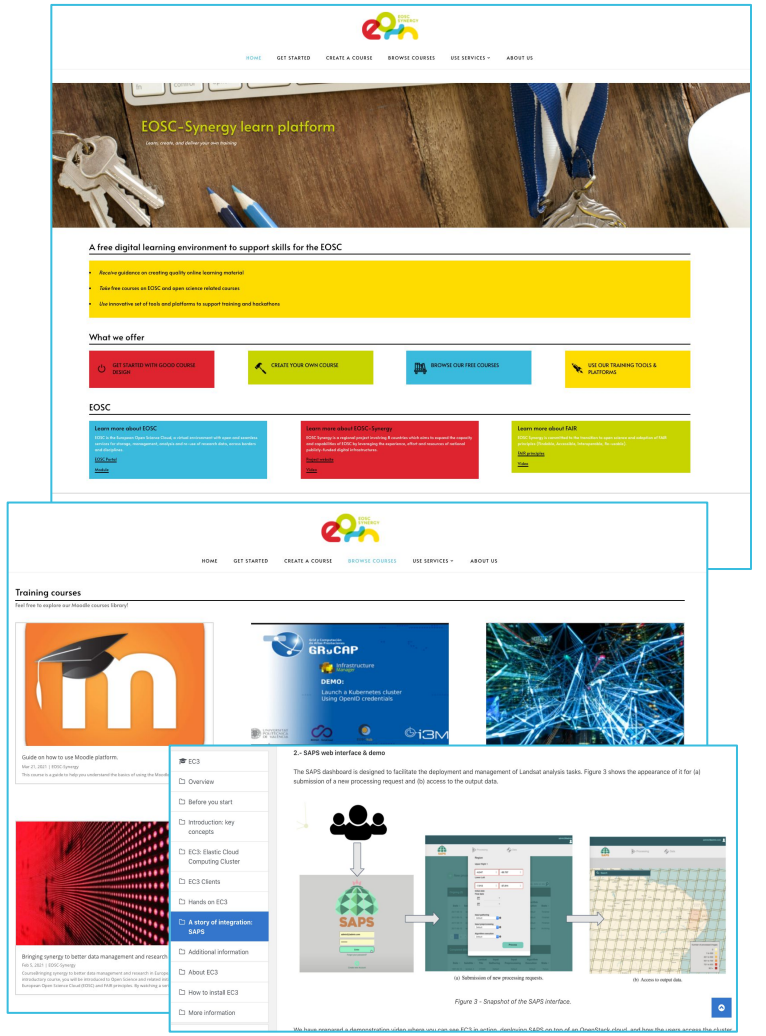
- Basic tutorials on EOSC
- Advanced tutorials regarding the EOSC-Synergy tools
- Domain specific tutorials regarding thematic services

Learn@Synergy and HaaS platforms

- Online platform for content creation/hosting of training material
- Service for running hackathons

Interaction with national education programmes

- Courses developed to be suitable for education programmes
- Had 10 different practical use cases of using the EOSC-Synergy platform or training materials during the activities of those institutions' educational processes in different countries.

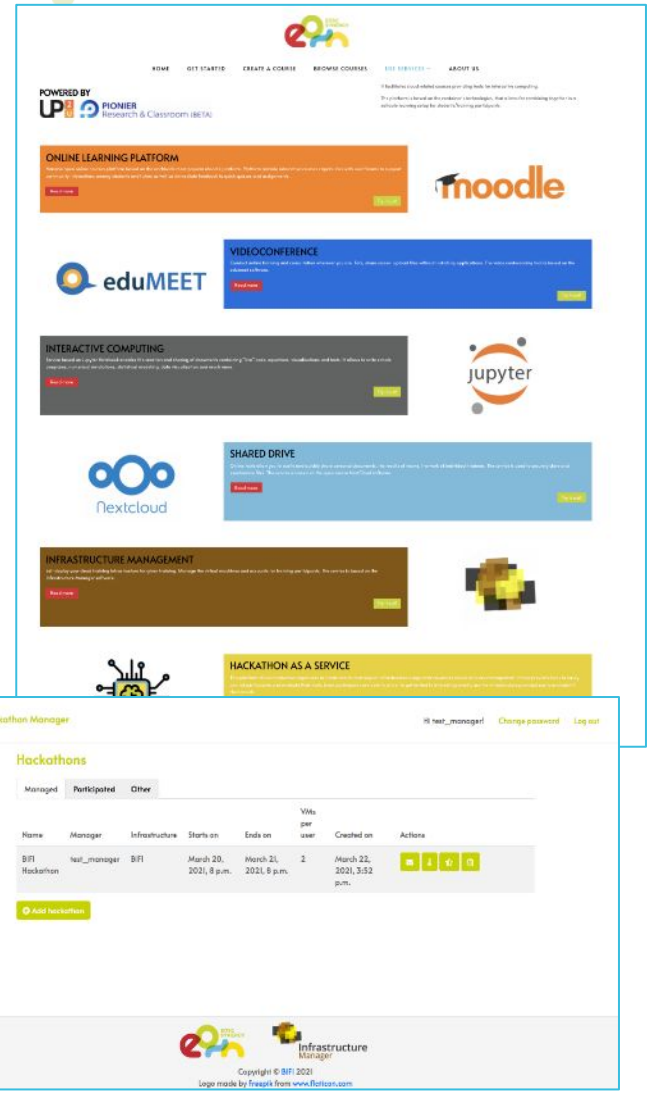


Learn@Synergy platform

An innovative online platform focussing on open tools

Modular set of tools for preparing and conducting tutorials
https://learn.eosc-synergy.eu/

- **MOOC:** customised **Moodle**
- **Videoconference**: based on **eduMEET**
- **Interactive computing**: based on **Jupyter Notebook**
- **Shared drive**: based on **NextCloud**
- **Infrastructure Manager** to create virtual machines and training accounts
- **Training material catalogue**
- **Hackathon-as-a-service** to facilitate the organisation of hackathons on the EOSC infrastructure and accessible through the EOSC Portal



Results

FAIR framework
Framework for validating EOSC FAIR data requirements. It provides Automated deployment of data repositories and fairness verification. One of the key components is the FAIR evaluator "FAIR EVA". FAIR EVA [...]

Read more

JePL
The Jenkins Pipeline Library (JePL) is a library to build and execute automated software quality verification pipelines using the Jenkins CI system. A new implementation of the INDIGO-DC jenkins-pipeline-library by [...]

Read more

SQAaaS Platform
A Software and Services Quality Assessment (SQA) for on-demand automated software validation, offered through the EOSC portal. The SQAaaS aims at contributing to the realization of the Open Science principles [...]

Read more

HaaS
The EOSC-Synergy Hackathon as a Service is a new approach to offer Infrastructure as a Service support for the execution of Hackatons. [...]

Read more

EOSC Training Platform
Our training platform is a set of tools, including procedures and best practices, for the creation and conduction EOSC related training courses. It facilitates cloud related courses providing tools for [...]

Read more

O3AS
The O3as service provides an invaluable tool to assist scientists in visualising ozone (O3) time series from extensive climate prediction model data and calculating and visualising dates for the recovery [...]

Read more

Handbook on EOSC infrastructure integration
How to make your infrastructure EOSC-aware [...]

Read more

Badge Scheme
The Quality Badge Scheme for software, services and FAIR data is a method to reward adherence to quality best practices for software and services. The badges are defined to motivate [...]

Services Quality Baseline
A set of quality baseline criteria for services based on best practices aiming at improving services quality. The key benefits are: Assess and assure the quality and maturity of services [...]

Software Quality Baseline
A set of quality baseline criteria for software based on good practices aiming at improving software quality. The purpose of this result is enhancing the visibility, accessibility and distribution of [...]

MSWSS
MSWSS is a service for modelling and analysis of Water Supply Systems which integrates the analysis of toxics in drinking-water supply networks with water distribution network simulation. MSWSS service will [...]

UMSA
UMSA is an untargeted mass-spectrometry analysis service from RECETOX (Research Centre for Toxic Compounds in the Environment at Masaryk University) in the Czech Republic. The service is evolving to a [...]

SDS-WAS
SDS-WAS (Barcelona Dust Regional Center from January 2022) is a service that aims at improving capabilities for more reliable sand and dust storm (SDS) forecasts. It supports institutional entities to [...]

www.eosc-synergy.eu

LAGO
The LAGO (Latin American Giant Observatory) Project is an extended astroparticle observatory at a global scale operated by the LAGO Collaboration, a non- centralized and distributed collaborative network of more [...]

Read more

OpenEBench
Used to evaluate Life Sciences research software, OpenEBench is an observatory for software quality based on the automated monitoring of FAIR for research software metrics and indicators. The OpenEBench platform [...]

Read more

SCIPION
Scipion is an application framework developed as a collaboration between many institutions including the Instruct Image Processing Center (I2PC) in Madrid to help the Structural Biology community to process Cryo [...]

Read more

SAPS
SAPS (SEB Automated Processing Service) is a service to estimate Evapotranspiration (ET) and other environmental data that can be applied, for example, on water management and the analysis of the [...]

Read more

G-Core
G-CORE is a production-ready technology used as a service at ESA's and national programs led by INDRA for the acquisition, storage, cataloguing and processing data from several Earth Observing (EO) [...]

Read more

WorSiCa
European-wide service for the detection of the coastlines changes, coastal inundation areas and inland waterbodies water detection [...]

Read more