# EOSC-hub

# The EOSC Early Adopter Programme

*A hands-on approach of EOSC for research communities*

**May 2020**

# The EOSC **Early Adopter** Programme

## *A hands-on approach of EOSC for research communities*

### What is the European Open Science Cloud?

The idea of a European Open Science Cloud (EOSC) took shape back in 2016, as a vision of the European Commission of a virtual environment for all researchers to store, manage, analyse and re-use data for research, innovation and educational purposes in Europe and beyond.

### What is the EOSC Early Adopter Programme?

The **EOSC Early Adopter Programme (EAP)** is an initiative launched by the EOSC-hub project aimed at research communities interested in exploring the latest state-of-art technologies and services offered by the European Open Science Cloud (EOSC).

The ultimate goal of the Early Adopter Programme is to gain insight into researchers' needs and, with their participation, drive the future developments of EOSC.

### Why the EOSC Early Adopter Programme?

→ To train and support researchers on the usage of the resources and services made available by EOSC-hub and partners via the EOSC marketplace

→ To gain insight into digital needs of researchers and their use cases

→ To foster a culture of co-operation between researchers and some of the EOSC service service providers

→ To foster confidence in the capability and capacity that will be provided via the EOSC

### About EOSC-hub

EOSC-hub is a three-year European project that brings together multiple service providers to create the Hub: a single contact point for European researchers and innovators to discover, access, use and reuse a broad spectrum of resources for advanced data-driven research.

EOSC-hub provides the expertise and resources to enable an active usage of the EOSC and to foster a culture of co-operation between researchers and EOSC providers.

This publication represents a collection of use cases – 13 in total – of the EOSC Early Adopter Programme, as selected by EOSC-hub.

## Goals

STARS4ALL is a European project raising awareness about the negative effects of artificial light on human wellbeing, biodiversity, visibility of stars, safety and energy waste.

This EAP use case aims at creating an EOSC light pollution community in order to use resources such as datasets, presentations, source code.

## Latest achievements

In collaboration with the EUDAT B2SHARE technical support team, the community defined the first version of the specific metadata including a mapping with existing metadata fields in B2SHARE. A mechanism to create a research object in B2SHARE has been agreed and the type of datasets (generated by the photometers) that are going to upload to B2SHARE have been identified. The implementation in B2SHARE is ongoing.

## EOSC-hub supporting services

B2Share, B2Find, PaN notebook, GEOSS web portal, EGI Cloud Compute, EGI Cloud Container Compute, EGI Notebooks

## Supported research community

STARS4ALL Foundation

## Partners

KNAW, EUDAT, EGI

# EMSO ERIC data management platform

## Goals

EMSO ERIC is a large-scale European research infrastructure for ocean observation.

This use case aims at transitioning the EMSO Data Management Platform (DMP) to pre-production and then to full production. The prototype DMP has been deployed within the EGI Federation and has been transitioning to production to the EOSC Marketplace as part of the EOSC-hub project.

## Latest achievements

The EMSO ERIC got access to the EGI Cloud Infrastructure and started the deployment of its data management platform. A number of EMSO services have been deployed and made available to the general public (EMSO Data Portal). Integration with EGI Check-in service is underway to enable the access via institutional credentials.

## EOSC-hub supporting services

EGI Check-in, Monitoring, Accounting, EGI Cloud Compute, EGI Online Storage, OpenAIRE AMNESIA

## Supported research community

EMSO ERIC

## Partners

EGI Foundation, CESGA, RECAS-BARI

# Mapping the sensitivity of mitigation scenarios to societal choices

**Science area:** Earth sciences, environmental engineering, economics

## Goals

This project aims to perform modeling studies to explore how future energy systems can evolve and to quantify the links between different aspects of the global energy systems in the context of international climate change policy and sustainable development.

These analyses utilise Integrated Assessment Models (IAMs), which are models of the energy, environment, and economic systems to quantify key variables of interest in these scenarios - such as emissions pathways consistent with international climate policy goals, tradeoffs of climate mitigation with land use and food security, among others.

This project will provide a proof-of-principle platform aimed at performing large scale analyses.

## Latest achievements

The community got access to the EGI Federated Cloud and is working to adapt their application to the cloud paradigm. An initial prototype is under preparation.

## EOSC-hub supporting services

EGI Check-in, EGI Cloud Compute, EGI Online Storage

## Supported research community

IIASA (The International Institute for Applied Systems Analysis)

## Partners

INFN

# Towards an e-infrastructure for plant phenotyping

**Science area:** Agricultural sciences

## Goals

In recent years, technological progress has been made in plant phenomics - major improvements concerning imaging and sensor technologies. High-throughput plant phenotyping platforms now produce massive datasets involving millions of plant images concerning hundreds of different genotypes at different phenological stages in both field and controlled environments. There is a need for an integrated and federated solution for data management and data processing. The open-source Phenotyping Hybrid Information System PHIS aims at organising these data and making them accessible and reusable to a larger scientific community.

## Latest achievements

A service level agreement has been agreed with EGI that guarantees the access to the EGI Cloud Compute infrastructure. The community is now working on deploying their application on top of the cloud resources and connecting them to their database of agriculture data. Integration with the EGI Check-in service has been planned to enable the access to the service via institutional credentials.

## EOSC-hub supporting services

EGI Cloud Compute, EGI Cloud Container Compute, B2SAFE, EGI DataHub, EGI Check-in, B2ACCESS, B2HANDLE, EGI Notebooks.

## Supported research community

French National Institute for Agricultural Research

## Partners

CINES

# Big Data analytics for agricultural monitoring using Copernicus sentinels and EU open data sets

**Science area:** Earth sciences, agriculture, environmental engineering

## Goals

The key aspect in this early adopter is to show how federated EOSC resources can facilitate a range of Sentinel data applications across agricultural user domains.

## Latest achievements

Access to EGI Cloud and CREODIAS (from CloudFerro) Infrastructure have been granted. A restful server has been set up to facilitate the access and the process of the Earth Observation data. A first exploratory work to use the EGI Notebooks is ongoing.

## EOSC-hub supporting services

EGI Cloud Compute, EGI Cloud Container Compute, CloudFerro Infrastructure, EODC SDIP infrastructure.

## Supported research community

European Commission, Joint Research Centre

## Partners

EODC, CloudFerro, CESNET, JRC, EGI Foundation

# Supporting FAIR data discoverability in clinical research: providing a global metadata repository (MDR) of clinical study object

**Science area:** Health sciences

## Goals

ECRIN is a non-profit organisation that links scientific partners and networks across Europe to facilitate multinational clinical research.

The main goal of this use case is to deploy the ECRIN MetaData Repository (MDR) CORE database and metadata conversion tool in the European Open Science Cloud.

Specific steps:

→ Setup of the EGI Data Hub service

→ Investigate the metadata scheme and requirements for future harvesting with B2FIND

→ Deploy the ECRIN Portal on the EGI Federated Cloud

→ Enable harvesting of a single global metadata repository (MDR) instance with B2FIND

→ Integration testing and enable harvesting of one or more MDR instance with B2FIND

## EOSC-hub supporting services

EGI Cloud Compute, EGI Cloud Container Compute, EGI DataHub, EUDAT B2FIND

## Supported research community

ECRIN (European Clinical Research Infrastructure Network)

## Partners

INFN, SURFSARA, DKRZ, CYFRONET

# Open AiiDA lab platform for cloud computing in materials science

**Science area:** Physical sciences

## Goals

This early adopter aims at providing access to Kubernetes managed infrastructure to support the deployment and operation of an open AiiDA lab instance in EOSC and supporting the authentication and authorisation of EOSC users into AiiDA lab.

## Specific steps

→ Integration with the EGI Check-in service

→ Deploy the AiiDA lab instance on the EGI Federated Cloud

→ Use of the AiiDA lab instance for a tutorial and support of at least 6 different simulation codes

→ On-boarding of users (>100 expected), testing and scaling, provide apps to connect to HPC resources

## EOSC-hub supporting services

EGI Check-in, EGI Cloud Compute, EGI Cloud Container Compute, EGI Online Storage, EUDAT B2HANDLE, OpenAire Zenodo, EOSC-hub monitoring, EGI IM/EC3.

## Supported research community

European researchers and industries via the MaX Centre of Excellence; European researchers and industries via H2020 MarketPlace project; Swiss researchers in the field of materials science via NCCR MARVEL; broader community of Swiss researchers via swissuniversities P-5 "Materials Cloud".

## Partners

GRNET, CESNET, UPV, OpenAire, MaX European Centre of Excellence, NCCR MARVEL, H2020 MarketPlace project, swissuniversities P-5 project "Materials Cloud".

**Science area:** Physical sciences
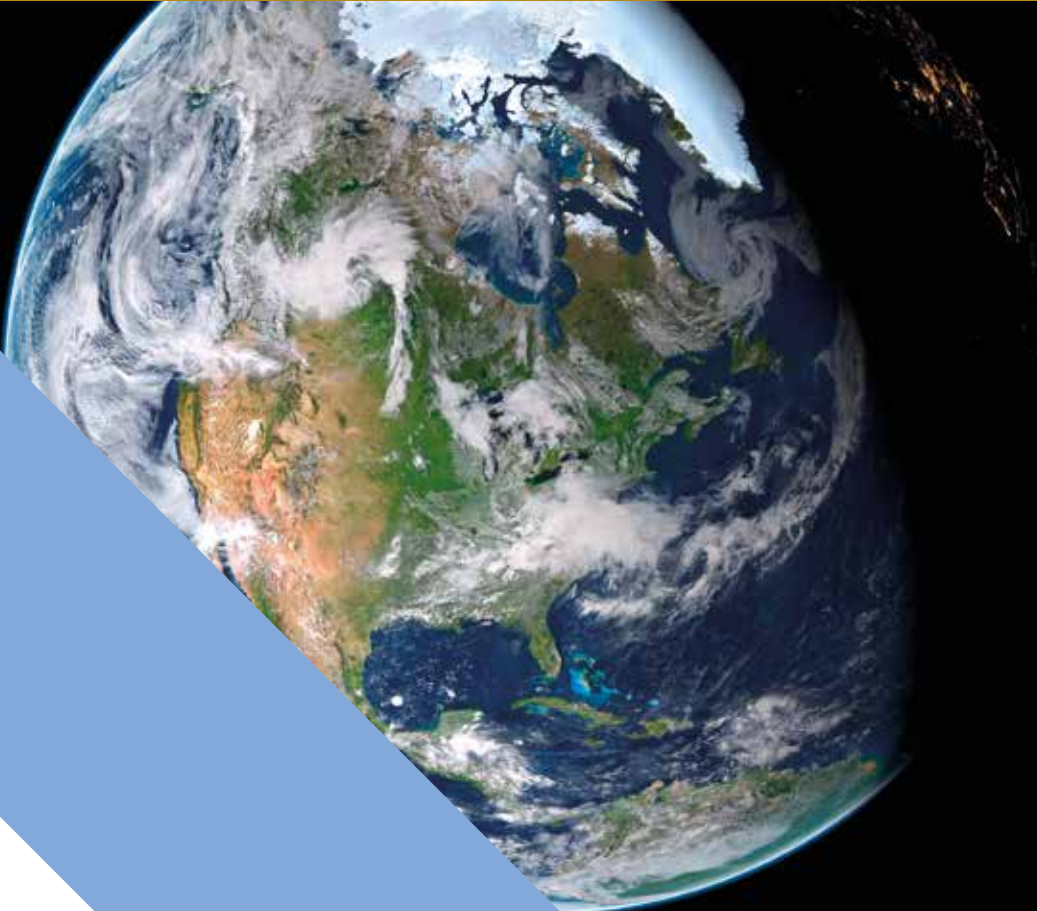
## Goals

VESPA aims at building a Virtual Observatory for Planetary Science, connecting all sorts of data in the field, allowing for rapid data search and providing modern tools to retrieve, cross-correlate, and display data and results of scientific analyses. The goal of the early adopter is to use the EOSC infrastructure to host the servers of the VESPA provider.

Specific steps:

→ Manage virtual machines (VMs) and containers in the EGI Federated Cloud and access storage

→ Automate VMs deployment and management, test harvesting of metadata with B2FIND service, configure eduTEAMS for AAI.

→ Document service deployment and have VMs template in AppDB, manage groups in eduTEAMS.

→ Onboard services in the EOSC Marketplace and explore sustainability options.

## EOSC-hub supporting services

EGI Cloud Compute, EGI Cloud Container Compute, EGI Check-in, EUDAT B2ACCESS, EUDAT B2FIND, EUDAT B2SAFE, GEANT eduTeams, EOSC-hub Monitoring, INDIGO PaaS, OpenAire Zenodo

## Supported research community

VESPA

## Partners

CESNET, GRNET, GEANT, Juliech, MPCD, OpenAire, DKRZ, IN2P3

# OpenBioMaps data management service for biological sciences and biodiversity conservation

**Science area:** Biology

## Goals

The OpenBioMaps (OBM) is a free and open-source database framework for scientific and conservation purposes used for data management by nature conservation institutes, biodiversity research and citizen science projects.

The goal of this early adopter is to deploy a service in EOSC that allows multiple users to run tasks for projects that collect nature conservation and biodiversity data.

Specific steps:

→ Integration with the EGI Cloud Compute service, deploy and configure OBM node in test environment.

→ Deploy the application automatically using INDIGO IM.

→ Deploy OBM node to production environment and assess EGI DataHub, B2DROP and B2FIND

→ Test the performance of the use case

## EOSC-hub supporting services

EGI Cloud Compute, EGI Online Storage, INDIGO IM, EGI DataHub, EUDAT B2DROP

## Supported research community

UNIDEB (University of Debrecen)

## Partners

IFCA, UPV, CYFRONET

# AGINFRA+: virtual research environments to support agriculture and food research communities

**Science area:** Agricultural sciences

## Goals

AGINFRA+ aims to exploit e-infrastructures to provide a sustainable channel addressing neighboring but not fully connected user communities around the agriculture and food domains.

A main goal of the early adopter is to create a DataMiner (DM) cluster and make it available to all the communities served by the D4Science infrastructure.

Specific steps:

→ Deploy the DM cluster in a single cloud provider and enable EOSC monitoring and accounting

→ Develop monitoring probes to check the status of the DM cluster

→ Integration of monitoring probes in the EOSC Monitoring service

→ Assessment of operational continuity of the application and onboarding of AGINFRA+ VREs in the EOSC Portal.

## EOSC-hub supporting services

EGI Check-in, EOSC-hub Monitoring, EGI Cloud Compute, EGI Online Storage

## Supported research communities:

AGINFRA+ project and CNR

## Partners

GRNET, STFC, IFCA

# EOSC DevOps framework and virtual infrastructure for ENVRI-FAIR common FAIR data services

**Science area:** Earth and related environmental sciences

## Goals

ENVRI-FAIR is a project that connects the Environmental Research Infrastructure (ENVRI) community to the European Open Science Cloud.

## Use cases of the early adopter:

→ Automated cloud execution for data workflow

→ Continuously testing and integration for ENVRI services

→ Notebook based environment for FAIR data access and processing

Specific steps:

→ Get resources and test the EOSC services

→ DevOps pipeline configured and demonstrated with at least 2 services

→ Automated workflow execution in cloud

→ Extended the involved service and explore sustainability options

## EOSC-hub supporting services

EGI Cloud Compute, EGI Online Storage, Jelastic PaaS, EGI Notebooks

## Supported research community

ENVRI-FAIR project

## Partners

CESGA, INFN, Jelastic

# Integration of toxicology and risk assessment services into the EOSC marketplace

**Science area:** Health sciences

## Goals

The OpenRiskNet project is developing an e-infrastructure for safety assessment, including toxicology and especially predictive toxicology, structural biology, bioinformatics, cheminformatics.

The main goal of the early adopter is to assess OpenRiskNet's services for EOSC integration.

Specific steps:

→ Testing of the EOSC services

→ Integration of EOSC AAI and deployment of OpenRiskNet services on EOSC cloud resources

→ Onboard and manage OpenRiskNet services in the EOSC Marketplace

## EOSC-hub supporting services

EGI Cloud Compute, EGI Online Storage, EOSC AAI

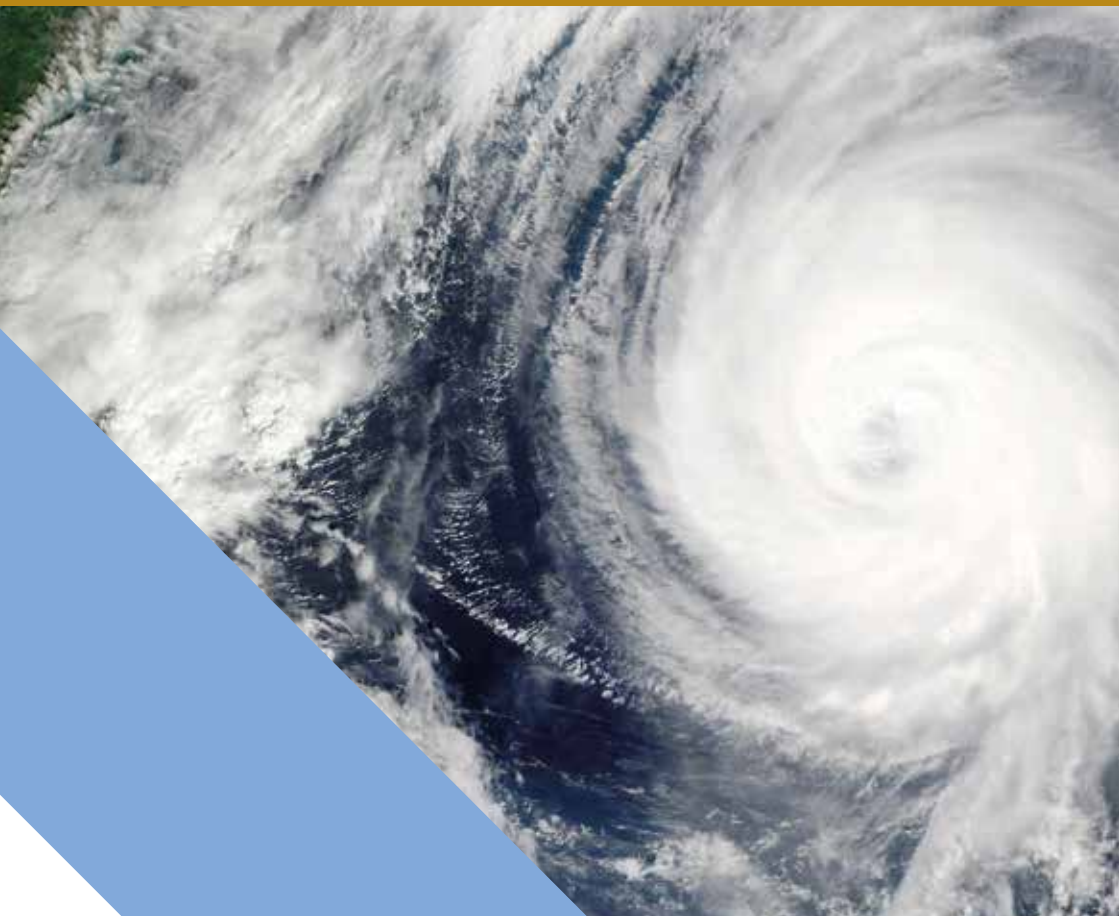## Supported research community

OpenRiskNet project

## Partners

INFN, UPV

# Towards a global federated framework for open science cloud

**Science area:** Environmental and health sciences

## Goals

This early adopter aims to allow researchers from Africa and China to use EOSC services on top of the Chinese Academy of Sciences (CNIC) resources.

3 use cases:

→ Disaster risk

→ Smart city

→ Precision medicine

Specific steps:

→ Access resources and migration on the CSTCloud

→ Integrate EGI Check-in service

→ Install and test the pilot use cases

→ Taiwan Typhoon forecasting with the WRF 4DVAR service, precision medicine using Elixir tools, Smart City application for storm surge and disaster damage assessment of ShenZhen

→ Use cases in production

## EOSC-hub supporting services

EGI Check-in, OCRE, OPENCoasT, DMCC, AGROS, WRF 4DVAR

## Supported research communities:

AASCTC (Sudan) and CNIC (China)

## Partners

GRNET, LNEC, ASGC, PSNC

eosc-hub.eu

@EOSC_eu

linkedin.com/company/eosc-hub

info@eosc-hub.eu