



Integrating a community

Earth sciences use cases

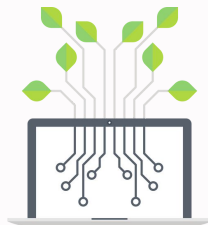
Samuel Keuchkerian - (CNRS)

Marie Jossé - Data Terra (CNRS)



FAIR-EASE

Building an interdomain digital architecture for distributed and integrated use of environmental data



FAIR-EASE Data Discovery
and Access Interdisciplinary
Service



FAIR-EASE Virtual
environments

FAIR-EASE Earth sciences use cases

5 pilots for an earth system model



Coastal Water Dynamics: focuses on the coastal marine environment near river estuaries, where important processes take place.



Earth Critical Zone: monitors land and soil degradation.



Volcano Space Observatory: monitors global volcanic activity, allowing the focus on any major volcanic eruption worldwide



Ocean Bio-Geochemical Observations:

addresses fundamental scientific questions regarding the health of marine ecosystems (e.g. ocean acidification, ...) and needs for ocean resource management.

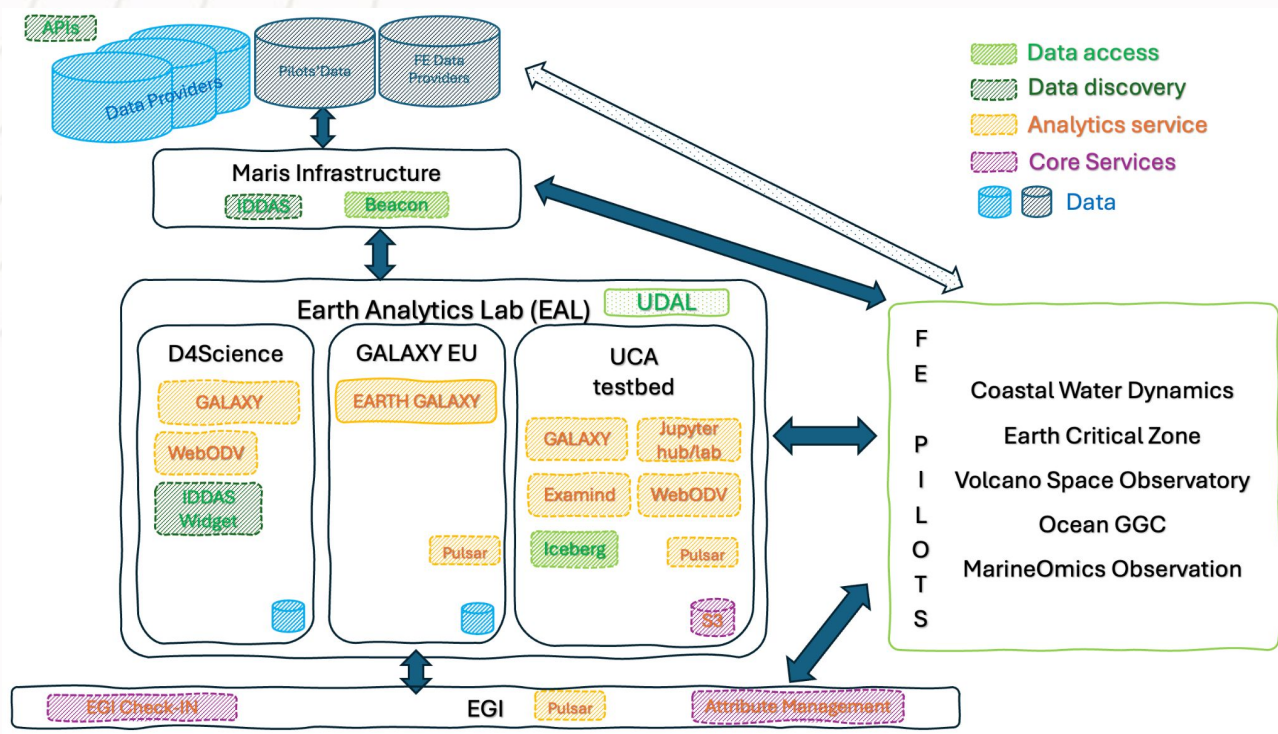


Marine Omics Observation:

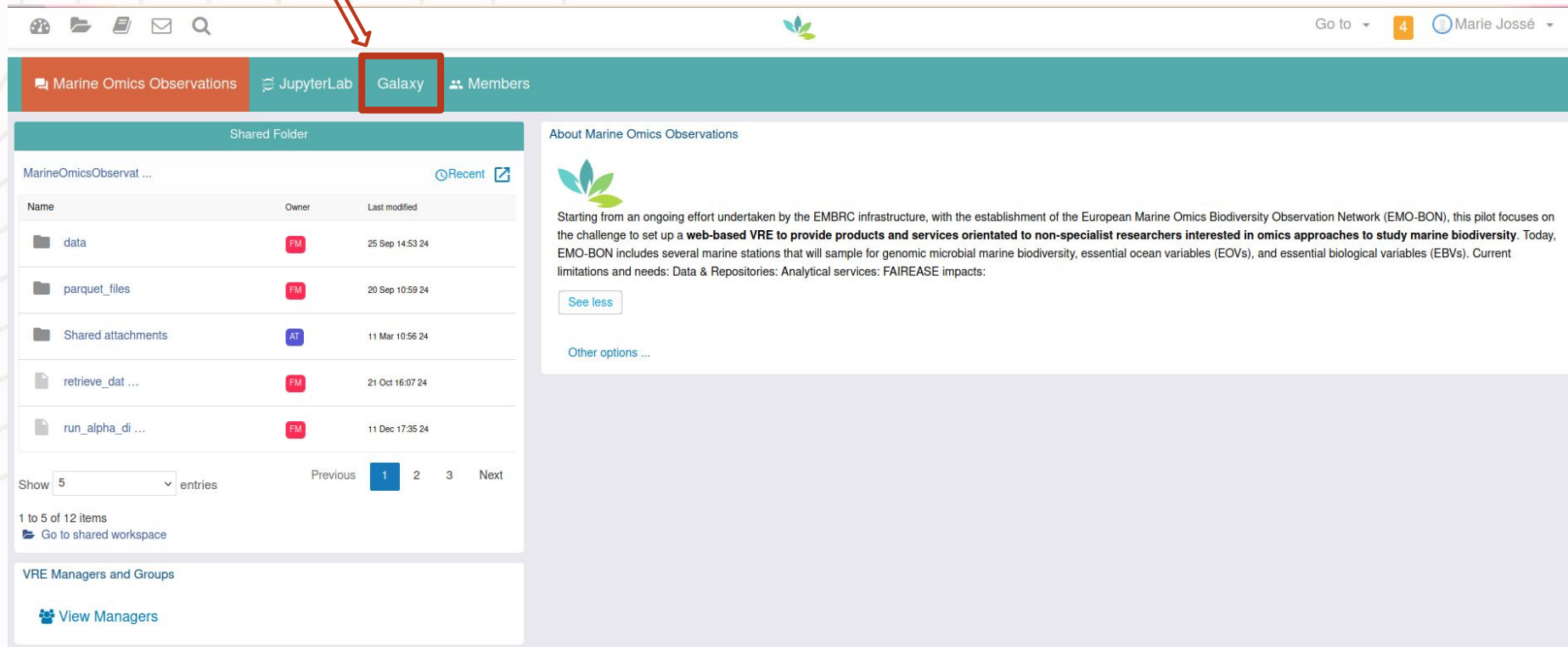
analyses of spatial- and time-comparable marine microbial metagenomics data sets for the exploration of biodiversity and its correlations with environmental quality



FAIR-EASE datalake infrastructure



FAIR-EASE Pilots on D4science V Labs



The screenshot displays the D4science V Labs interface. At the top, a navigation bar contains icons for home, files, documents, mail, and search. On the right, there is a 'Go to' dropdown, a notification badge with the number '4', and a user profile for 'Marie Jossé'. Below this, a teal navigation bar features four tabs: 'Marine Omics Observations', 'JupyterLab', 'Galaxy' (which is highlighted with a red box and pointed to by a red arrow), and 'Members'. The main content area is divided into two panels. The left panel, titled 'Shared Folder', shows a file browser for 'MarineOmicsObservat ...'. It contains a table with columns for 'Name', 'Owner', and 'Last modified'. The table lists five items: 'data' (owner FM, modified 25 Sep 14:53 24), 'parquet_files' (owner FM, modified 20 Sep 10:59 24), 'Shared attachments' (owner AT, modified 11 Mar 10:56 24), 'retrieve_dat ...' (owner FM, modified 21 Oct 16:07 24), and 'run_alpha_di ...' (owner FM, modified 11 Dec 17:35 24). Below the table, there is a 'Show 5 entries' dropdown and pagination links for 'Previous', '1' (selected), '2', '3', and 'Next'. A note at the bottom of this panel states '1 to 5 of 12 items' and provides a 'Go to shared workspace' link. The right panel, titled 'About Marine Omics Observations', features the EMO-BON logo and a paragraph of text: 'Starting from an ongoing effort undertaken by the EMBRC infrastructure, with the establishment of the European Marine Omics Biodiversity Observation Network (EMO-BON), this pilot focuses on the challenge to set up a **web-based VRE to provide products and services orientated to non-specialist researchers interested in omics approaches to study marine biodiversity**. Today, EMO-BON includes several marine stations that will sample for genomic microbial marine biodiversity, essential ocean variables (EOVs), and essential biological variables (EBVs). Current limitations and needs: Data & Repositories: Analytical services: FAIREASE impacts:'. Below the text are links for 'See less' and 'Other options ...'. At the bottom left of the interface, there is a section titled 'VRE Managers and Groups' with a 'View Managers' link.

Go to 4 Marie Jossé

Marine Omics Observations JupyterLab **Galaxy** Members

Shared Folder

MarineOmicsObservat ... Recent

Name	Owner	Last modified
data	FM	25 Sep 14:53 24
parquet_files	FM	20 Sep 10:59 24
Shared attachments	AT	11 Mar 10:56 24
retrieve_dat ...	FM	21 Oct 16:07 24
run_alpha_di ...	FM	11 Dec 17:35 24

Show 5 entries Previous 1 2 3 Next

1 to 5 of 12 items
Go to shared workspace

VRE Managers and Groups
View Managers

About Marine Omics Observations

Starting from an ongoing effort undertaken by the EMBRC infrastructure, with the establishment of the European Marine Omics Biodiversity Observation Network (EMO-BON), this pilot focuses on the challenge to set up a **web-based VRE to provide products and services orientated to non-specialist researchers interested in omics approaches to study marine biodiversity**. Today, EMO-BON includes several marine stations that will sample for genomic microbial marine biodiversity, essential ocean variables (EOVs), and essential biological variables (EBVs). Current limitations and needs: Data & Repositories: Analytical services: FAIREASE impacts:

See less

Other options ...

← → 🔍 https://earth-system.usegalaxy.eu

Import bookmarks... OBM - Agenda The Workflow Run RO... Galaxy | Europe conda-forge | commu... GCC CoFest 2024 - Cro... Galaxy | Configured b... Galaxy Home - Galaxy Comm... Boîte de réception (3...

Galaxy Earth System

Workflow Visualize Data Help User

Using 21%

Upload

Tools

Workflows

Workflow Invocations

Visualization

Histories

History Multiview

Datasets

Pages

Notifications

Settings

Tools

⌵ ☆

Get Data

Send Data

Collection Operations

GENERAL TEXT TOOLS

Text Manipulation

Filter and Sort

Join, Subtract and Group

Convert Formats

Annotation

Climate Analysis

GENERAL INTERACTIVE TOOLS

Interactive tools

EARTH AND ENVIRONMENTAL DYNAMICS

Water Coastal Dynamics

Earth Critical Zone

Volcano

ENVIRONMENTAL BIO-GEOCHEMICAL ASSETS

Bio-geochemical

BIODIVERSITY OBSERVATION

Marine Omics

Biodiversity data exploration

WORKFLOWS

All workflows

FAIR-EASE

Welcome to Galaxy for Earth System and environment

Galaxy for Earth System and environment was implemented within the project [Fair-Ease](#). It's a virtual platform to process, analyse and visualize Earth System, Environment and Biodiversity data. It is based on the [Galaxy framework](#), which guarantees simple access, easy extension, flexible adaption to personal and security needs, and sophisticated analyses independent of command-line knowledge.

Content

1. [Welcome to Galaxy for Earth System and environment](#)
2. [Get started](#)
3. [Training](#)
4. [Tools](#)
5. [Workflows](#)
6. [References](#)

Get started

Are you new to Galaxy for Earth System sciences, or returning after a long time, and looking for help to get started? Take a [guided tour](#) through a detailed tutorial on the Galaxy Training Network.

Training

We are working in close collaboration with the [Galaxy Training Network \(GTN\)](#) to develop training materials of data analyses based on Galaxy. If you

History

search datasets

Argo BGC Pangeo

ocean

2.14 MB

✓

25: FLUX on text prompt

22: version

21: TIME

20: N_POINTS

19: LONGITUDE

18: LATITUDE

16: Resulting map bgc

a list with 1 png dataset

15: Netcdf bgc

a list with 0 datasets

12: Timeseries plot bgc

11: Timeseries bgc tabular

8: info bgc

7: bgc metadata tabular

6: Coordinate info bgc

a list with 5 tabular datasets

2: Argo BGC data



Galaxy Training Network

An easy way to learn how to use Galaxy and improve your skills on various domains for instance a set of tutorials are available on FAIR management

https://training.galaxyproject.org/training-material/

Welcome to Galaxy Training!
Collection of tutorials developed and maintained by the worldwide Galaxy community

Galaxy for Scientists
We have separated the tutorials into several categories based on field and technology. We are exploring other ways to organise the tutorials going forward!

Start Here

Topic	Tutorials
Introduction to Galaxy Analyses	13
Using Galaxy and Managing your Data	22

Not sure where to start?
Try the NGS Basics Learning Path! [Start Learning](#)

Scientific Fields

Topic	Tutorials
Climate	12
Computational chemistry	9

Quickstart

- Learning Pathways
- Galaxy for SysAdmins
- Galaxy for Developers
- Galaxy for Teachers

Upcoming Events
Check out upcoming events around the Galaxy!

- January 26, 2025
Galaxy at SURF Research Cloud workshop
- February 4 - 6, 2025
Code & Collaborate: The FAIRtyale of Software Development

- A catalog of tutorials
- Pathways on a dedicated topic
- Classes, courses, webinars, and other interactive events

A community for the community
<https://training.galaxyproject.org/training-material/>



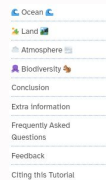
Galaxy Training Network

FAIR-EASE contribution (for earth sciences)

Earth sciences discovery tutorials

Thematic tutorials on ocean, land, atmosphere, and biosphere

Development tutorials and pathways to build a subdomain and a community with Galaxy



This tutorial aims at familiarizing you with Earth Science and discovering the earth data available on Galaxy. The target audience is not a scientist but anyone interested in learning about Earth system.

Agenda

In this tutorial, we will cover:

- Ocean**
 - The Argo program
 - Copernicus Data Space Ecosystem
 - EMODnet Chemistry
 - Copernicus Marine Data Store
- Land**
 - Copernicus Data Space Ecosystem
 - GGIS (Geographical Information System)
- Atmosphere**
 - Copernicus Data Space Ecosystem
 - Climate Data Store
- Biodiversity**
 - Marine biodiversity
 - Land biodiversity
- Conclusion
- Extra Information

Getting your hands-on earth data

Author(s): Marie Josse

Reviewers: [User Avatars]

Ocean's variables study

Author(s): Marie Josse

Reviewers: [User Avatars]

Overview

Questions:

- How to process extract ocean's variables?
- How to use ODV collections?
- How to create climatological estimates?

Objectives:

- Deals with ODV collection with data originating from Emodnet chemistry
- Visualise ocean variables to study climate changes

Tool development for a nice & shiny subdomain

Discover Galaxy's communities and learn how to create your subdomain and enrich it by writing, testing and submitting your tools on Galaxy. This learning pathway will guide you through all the steps required to build a tool for Galaxy with Planemo for batch tools and how write an interactive tool.

subdomain community tool development

3-day course dev

Importance of the team in the sustainability of Galaxy

A vibrant team of :



Administrators



Developpers



Trainers

**BUT with a strong focus on life
sciences**



**Need the same team for earth
sciences**

Freiburg Team



Björn Grüning

Dr. rer. nat., Researcher, Head of the Team

✉ gruening@informatik.uni-freiburg.de
 📞 bgruening
 ☎ +49(0) 761 - 203 54130
 📍 Build.: 079, Room: -1006
 🗣 [m] @ in 📺 📻 📡



Bérénice Batut

Ph.D., Researcher

✉ bebatut@informatik.uni-freiburg.de
 📞 bebatut
 ☎ +49(0) 761 - 203 54126
 🌐 <http://research.bebatut.fr>
 🗣 @ in 📺 📻 📡



Anika Erxleben-Eggenhofer

Dr. rer. nat., Researcher

✉ erxleben@informatik.uni-freiburg.de
 📞 erxleben
 ☎ +49(0) 761 - 203 54130
 📍 Build.: 079, Room: -1006
 🗣 [m] in 📺



Wolfgang Maier

Dr. phil. nat., Researcher

✉ maierw@informatik.uni-freiburg.de
 📞 wm75
 ☎ +49(0) 761 - 203 54126



Helena Rasche

B.Sc. Biochem., Technician

✉ hxr@informatik.uni-freiburg.de
 📞 hexylena
 🗣 Galaxy Administrator

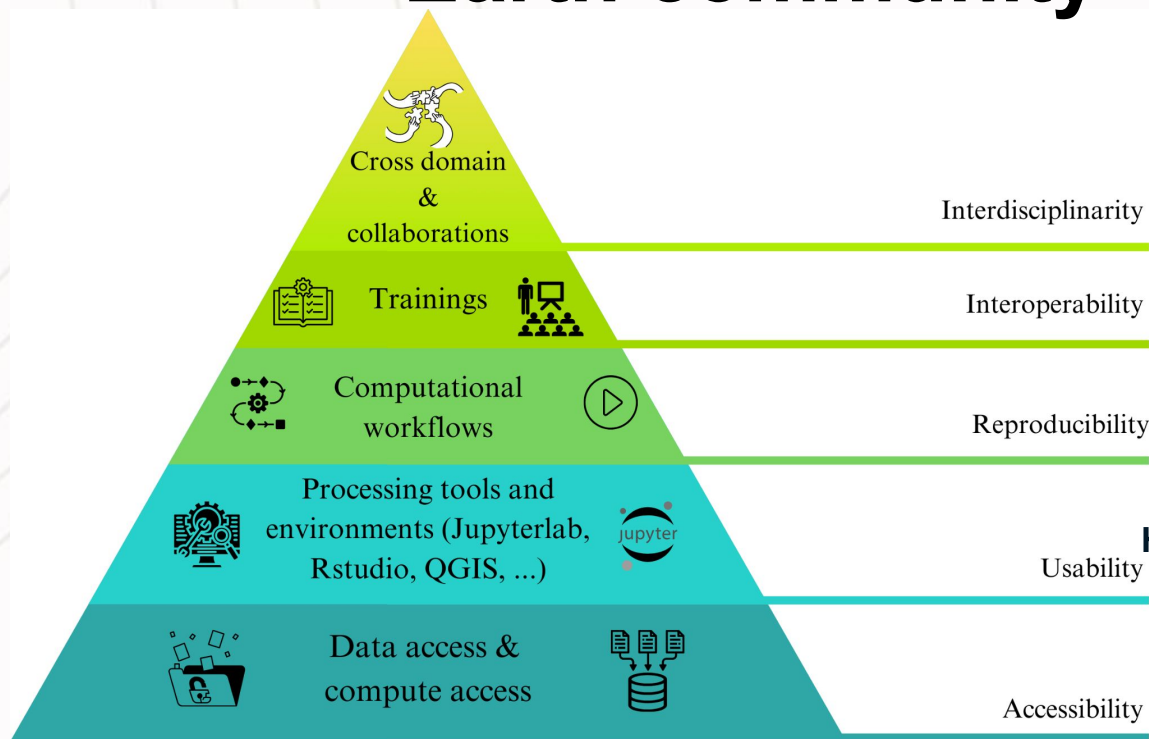


Paul Zierep

Dr. rer. nat., Researcher

✉ zierep@informatik.uni-freiburg.de
 📞 paulzierep
 ☎ +49(0) 761 - 203 54130

Earth community



How to maintain it

How we built it

Services for EOSC in the proposal for a node (and the EOSC EU node)

French national digital infrastructures

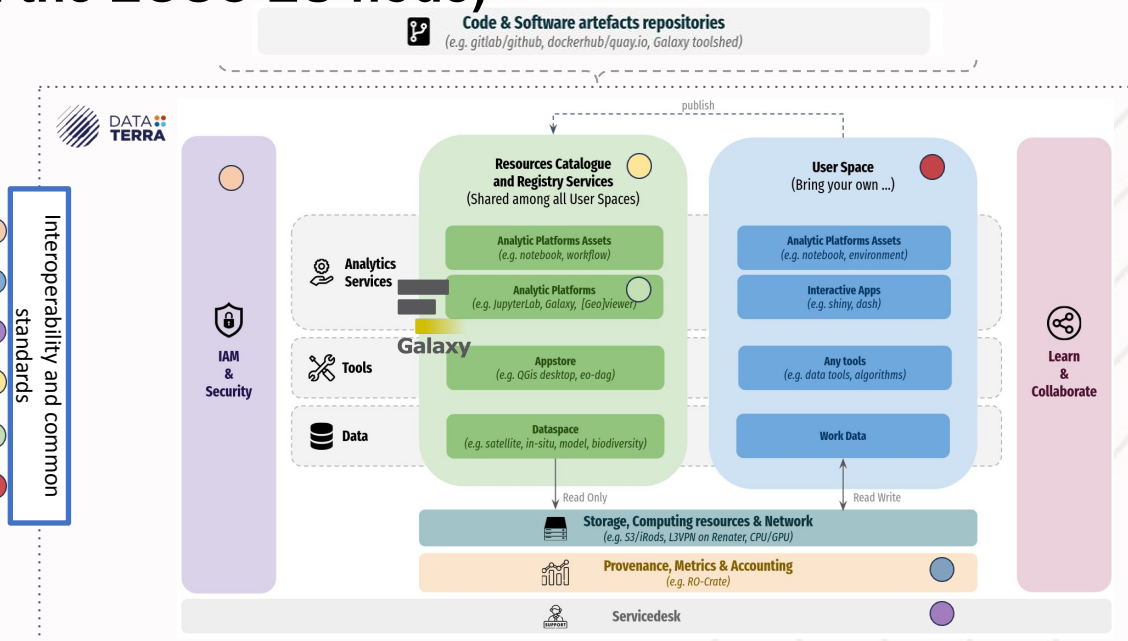
- Renater (Geant);
- France Grille (EGI), & Mesonet > GENCI / EuroHPC: IDRIS, CINES, TGCC;
- National & regional labelled data centres and meso-centres.

EOSC Federation

- D-T data and services accessible through the EOSC EU node;
- Services interoperability with thematic cluster nodes and related national nodes.

Core services

- Distributed data storage and management;
- Large data transfer (files, objects);
- User spaces (interactive notebooks, virtual machine, container images);
- HPC/Cloud computing services
- Federated AAI.



Gaia data timeline allows to co-develop with EOSC EU node and third party (EOSC Federation)

Data Terra services designed to support EOSC EU Node and third party core services

EOSC Federation (EOSC EU node and other nodes)

- Identity Management
- Monitoring and Accounting
- Service Management System
- Resources Catalogues and Registry Services
- Application Workflow Management
- User Space

Integrated Earth System Observation - Data Terra

Research Infrastructure to access, process and combine data



MAIN

Develop a global system for accessing and processing observation data (satellite, in situ), value-added products and services to

MISSION

observe, understand and predict in an integrated manner the functioning and evolution of the Earth system.



Partners

French scientific organizations and universities



5 thematic hubs

Atmosphere



Ocean



Solid earth



Land Surfaces



Biodiversity



Pôle National de Données de Biodiversité

Cross-cutting device

Satellite images



A multidisciplinary approach because it calls on work in several areas of Earth System sciences

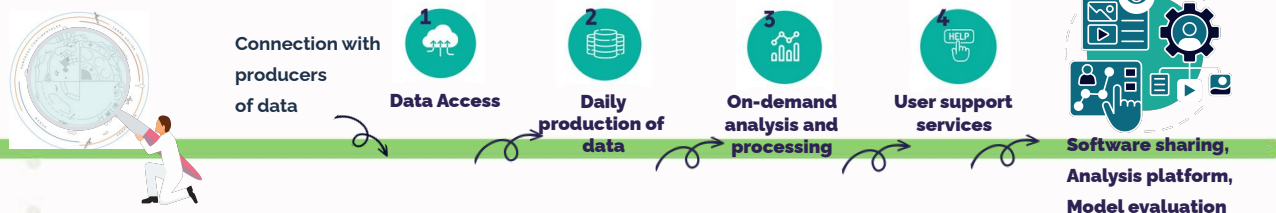
An inclusive project that goes beyond the scientific circle with an approach also oriented towards field actors and participatory data

OUR

SERVICES

Facilitating the cross-referencing of observations and the modeling of Earth System data

The IR Data Terra offers services around Earth system observation data. The objective is to provide interoperable and interdisciplinary services at all levels.



wednesday session



Data storage and handling

Samuel Keuchkerian - (CNRS/FAIR-EASE)

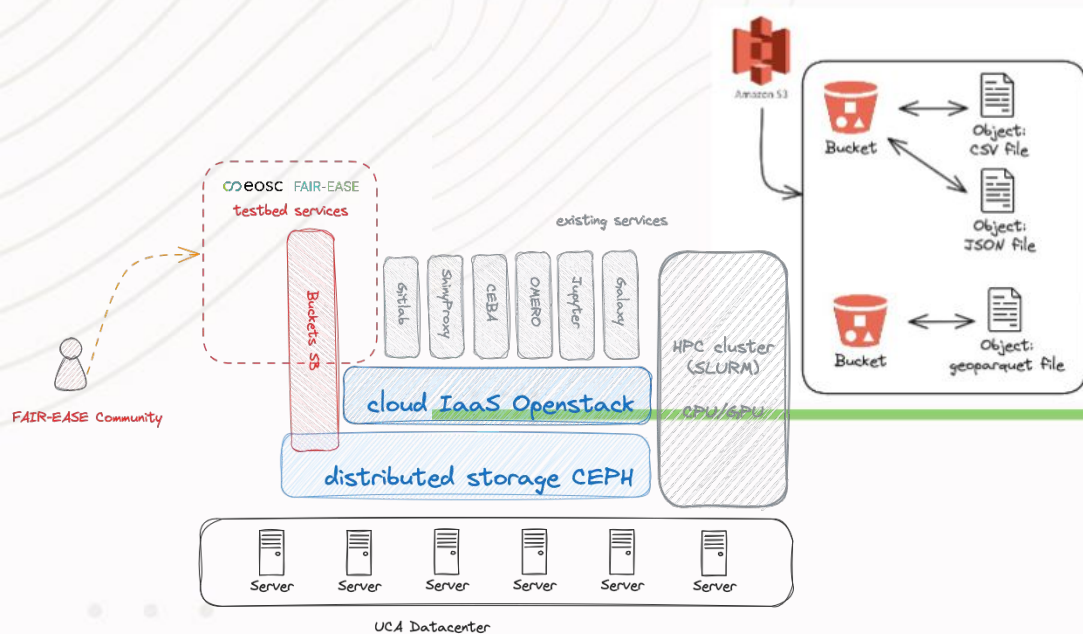
Marie Jossé - Data Terra (CNRS)



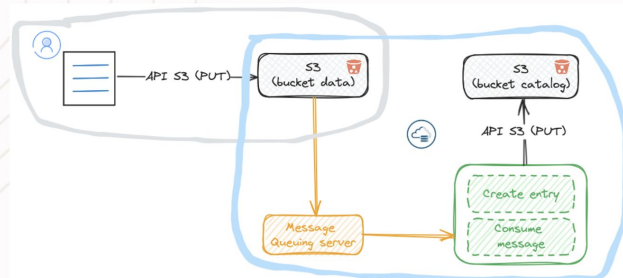
Integrating New Data Storage and Access Paradigms in pilot's practices

"S3 Testbed Tests" Antoine Mahul, David Sarramia, Damian Smyth

S3 concepts reminder



Testing interactions with a static STAC

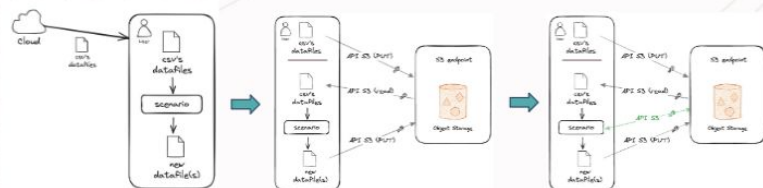


Can we manipulate a static STAC catalog directly on S3?

Can we use S3 notifications to automatically update a static catalog?

Can we manage metadata automatically from S3 to STAC?

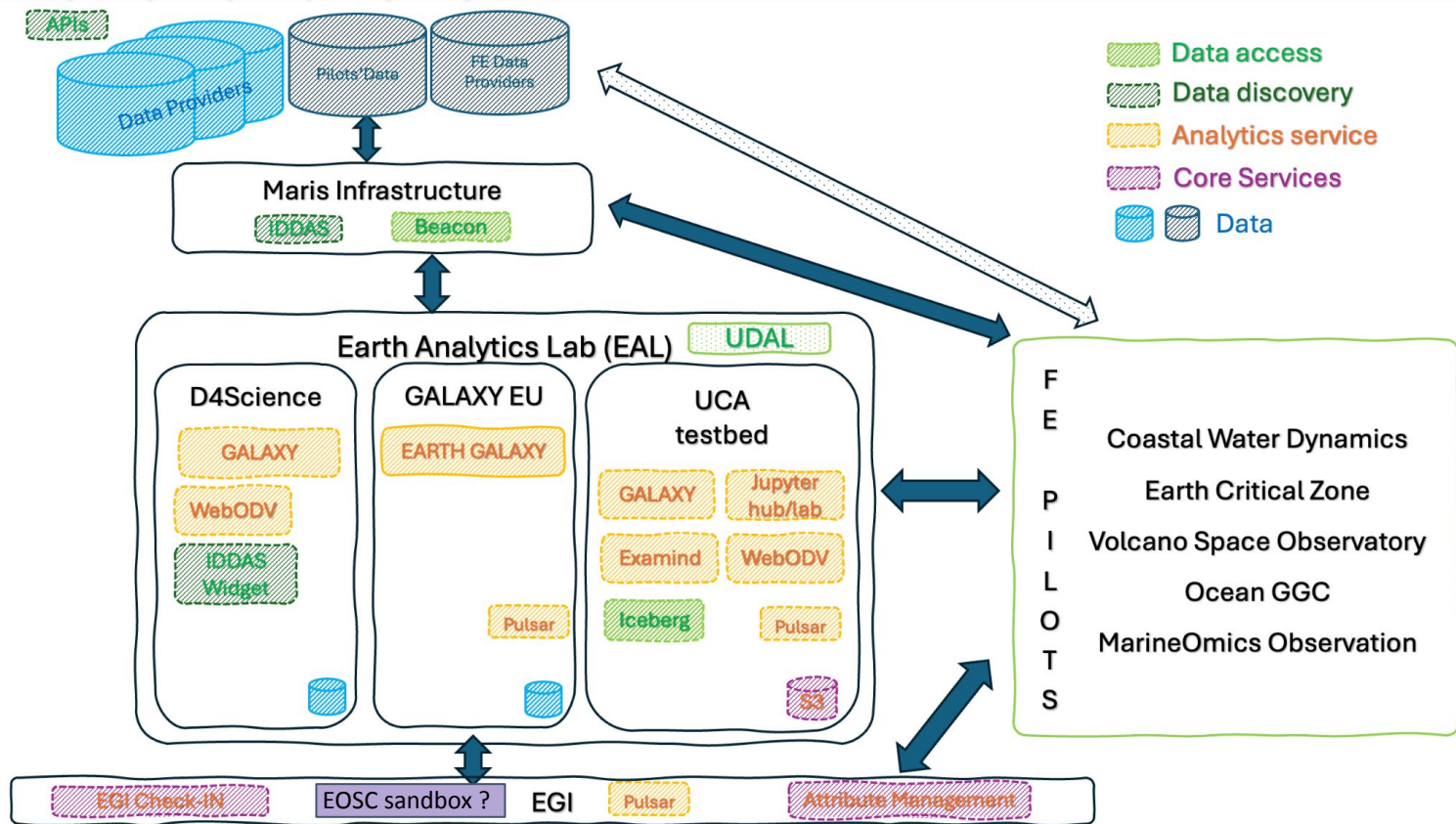
Testing interactions between jupyter notebook, third party softwares and S3



Common practice

Using S3 for input and output

Using S3 as a local storage





Adding file-sources to Galaxy



Under Development!

This tutorial is not in its final state. The content may change a lot in the next months. Because of this status, it is also not listed in the topic pages.



Marie Josse

Overview

Questions:

- How to set up an S3 bucket

Objectives:

- Add your S3 bucket on Galaxy

Time estimation: 15 minutes

Supporting Materials:

Published: Jan 20, 2025

Last modification: Jan 20, 2025

License: Tutorial Content is licensed under [Creative Commons Attribution 4.0 International License](https://creativecommons.org/licenses/by/4.0/). The GTN Framework is licensed under [MIT](https://mit.edu/)

Revision: 1


Find the information you
need

Add the S3 bucket

Conclusion

This tutorial demonstrates how to implement an S3 bucket as a Galaxy file-source within Galaxy. We will add here the public Argo data Amazon S3 bucket. Argo is an international program that observes the interior of the ocean with a fleet of profiling floats drifting in the deep ocean currents (<https://argo.ucsd.edu>). It started 20 years ago and is a dataset of 5 billion in situ ocean observations from 18.000 profiling floats (4.000 active). The Argo GDAC dataset is a collection of 18.000 NetCDF files. It is a major asset for ocean and climate science and a contributor to IOCCP reports.

Agenda

In this tutorial, we will cover:

1. [Find the information you need](#)
2. [Add the S3 bucket](#)
3. [Conclusion](#)

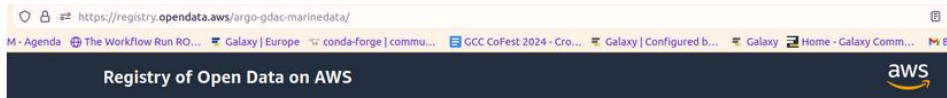
Find the information you need

Hands-on: Find an S3 bucket

Go on [Amazon Sustainability Data Initiative](#).

There you can visit the [catalog of data](#), and by searching for Argo you can directly get to the [Argo registry](#).

On this last page you'll find all the information you'll need to add the S3 bucket to Galaxy



Description

Argo is an international program to observe the interior of the ocean with a fleet of profiling floats drifting in the deep ocean currents (<https://argo.ucsd.edu>). Argo GDAC is a dataset of 5 billion in situ ocean observations from 18,000 profiling floats (4,000 active) which started 20 years ago. Argo GDAC dataset is a collection of 18,000 NetCDF files. It is a major asset for ocean and climate science, a contributor to IOCCP reports.

Update Frequency

Data is updated daily.

License

Open data, there are no restrictions on the use of this data. <https://creativecommons.org/licenses/by/4.0/>

Documentation

<http://www.argodatamgt.org/Documentation>

Managed By

Euro-Argo

See all datasets managed by [Euro-Argo](#).

Contact

Resources on AWS

Description

Argo GDAC data and metadata

Resource type

S3 Bucket

Amazon Resource Name (ARN)

`arn:aws:s3:::argo-gdac-sandbox`

AWS Region

`eu-west-3`

AWS CLI Access (No AWS account required)

`aws s3 ls --no-sign-request s3://argo-gdac-sandbox/`

Explore

[Browse Bucket](#)



Add the S3 bucket

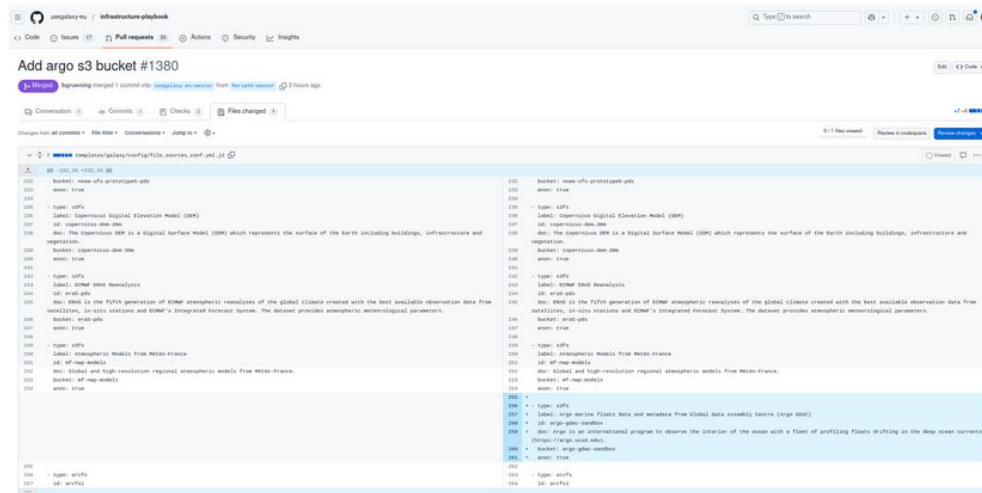


Hands-on: Add on Galaxy

- If not already done clone the Galaxy Europe [Infrastructure-playbook repo](#)
- Create a branch on your fork
- Go to the file **file_sources_conf.yml** in **templates/galaxy/config/**

There you can edit the file and add your S3 bucket by adding a Argo specific section, like in the following:

```
- type: s3fs
label: Argo marine floats data and metadata from Global Data Assembly Centre (Argo GDAC)
id: argo-gdac-sandbox
doc: Argo is an international program to observe the interior of the ocean with a fleet of profiling floats
bucket: argo-gdac-sandbox
anon: true
```



Finally, commit your changes and write a nice message for the admin when you open your Pull Request.

Upload

Tools

Workflows

Workflow Invocations

Visualization

Histories

History Multiview

Datasets

Pages

Tools

Get Data

Send Data

Collection Operations

GENERAL TEXT TOOLS

Text Manipulation

Filter and Sort

Join, Subtract and Group

Convert Formats

Annotation

Climate Analysis

GENERAL INTERACTIVE TOOLS

Interactive tools

Earth System and environment

Water Coastal Dynamics

Earth Critical Zone

Volcano

Upload from

Regular

Argo

Label

Argo marine floats data and metadata from Global Data Assembly Centre (Argo GDAC)

Details

Argo is an international program to observe the interior of the ocean with a fleet of profiling floats drifting in the deep ocean currents (<https://argo.ucsd.edu>).

«

1

2

»

Cancel

Ok

History

search datasets

Argo BGC Pangeo

ocean

2.14 MB

14

9

2

25: FLUX on text prompt

22: version

21: TIME

20: N_POINTS

19: LONGITUDE

18: LATITUDE

16: Resulting map bgc

a list with 1 png dataset

15: Netcdf bgc

a list with 0 datasets

[← Back to News](#)

Bring Your Own Storage with Galaxy

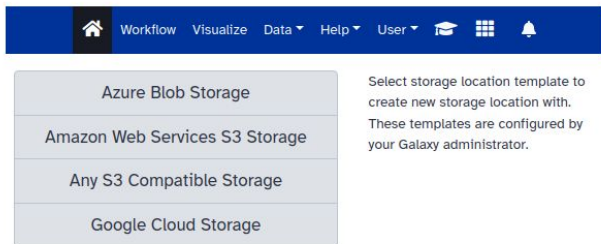
Connect Galaxy with external storage resources easily

September 20, 2024

Introduction

The [EuroScienceGateway Project](#) is streamlining the way that users *Bring their Own Storage (BYOS)* to Galaxy. This post covers a specific case in which a user has access to storage resources in the [EGI Federated Cloud](#) and wants to connect it to the [EU Galaxy instance](#). However, similar steps can be followed to connect a different Galaxy instance with storage capacity using [Azure Blob Storage](#), [Amazon Web Services S3 Storage](#), and [Google Cloud Storage](#).

In order to connect [Galaxy](#) with external cloud storage, the user needs to click on the [User](#) drop-down menu, select [Preferences](#) and then [Manage Your Storage Locations](#). This will display a list of the cloud storage services already connected to Galaxy. The first time it will be empty and the user can connect to a new cloud storage service clicking on [Create](#). The following options are given to the user:



Generally speaking the user needs the following information to connect Galaxy with cloud storage:

1. Name of the bucket
2. Access key
3. Secret key

Getting S3 storage from EGI

Users with access to cloud resources in the [EGI Federated Cloud](#) can benefit from the available [TOSCA template](#) in [Infrastructure Manager \(or IM\)](#) for the automated deployment of [MinIO](#).

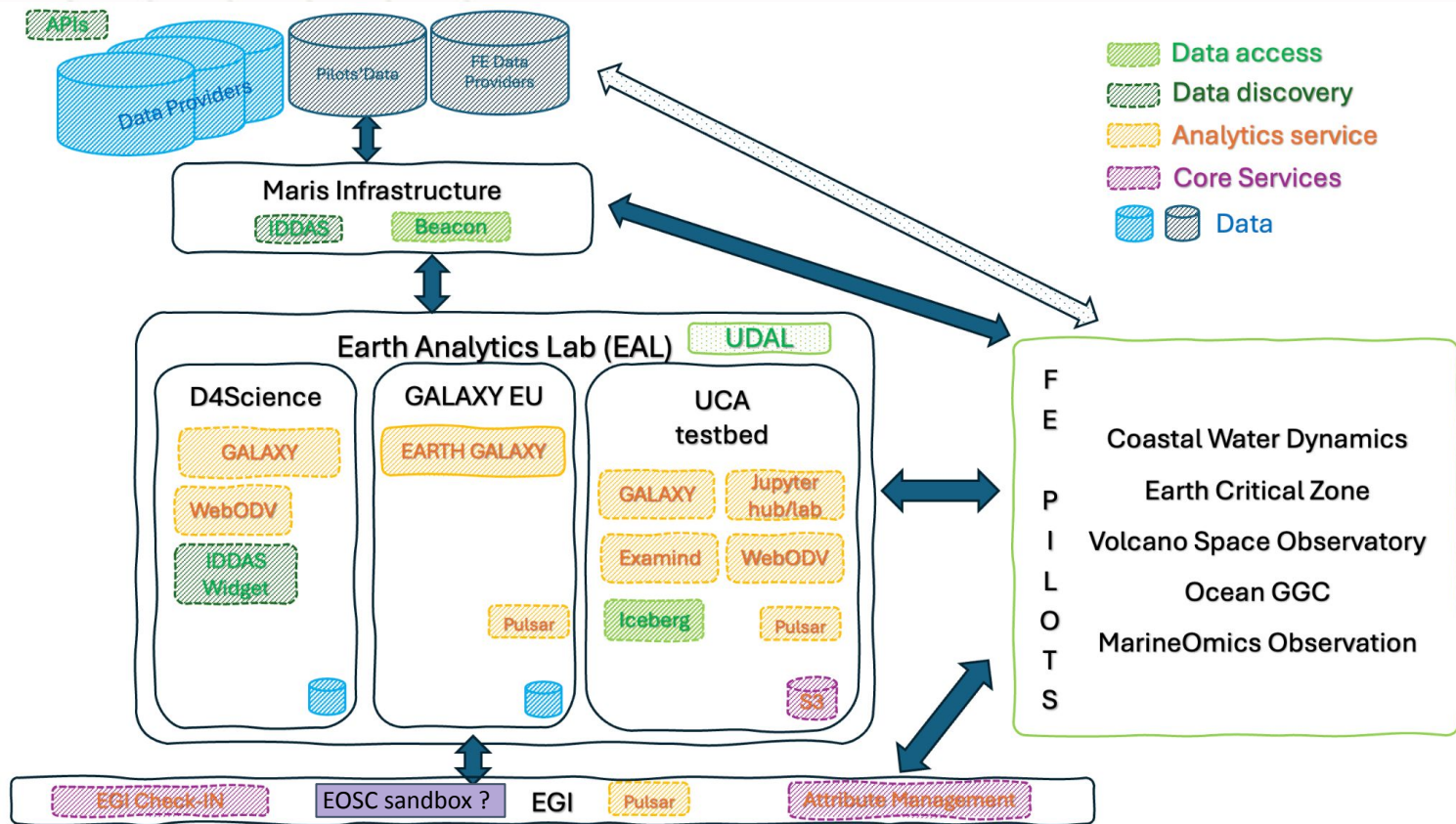
Below are the steps that a user needs to follow to make use of computing resources in the EGI Federated Cloud:

1. Create an [EGI Check-in account](#).
2. Enroll in a [Virtual Organization \(VO\)](#). You need to wait for approval before moving on.
3. Once you are a member of a VO, [configure credentials in Infrastructure Manager](#).
4. In [Infrastructure Manager](#) select [Deploy a VM](#) and click [Configure](#).
5. Select [MinIO](#) and click [Add](#).
6. Configure deployment details for MinIO and click [Submit](#).

A prerequisite for the deployment of MinIO with IM is the use of some sort of Dynamic DNS service. EGI also offers a [Dynamic DNS service](#) for free for users with an EGI Check-in account (see a [tutorial](#) for further help). Using a Dynamic DNS service the user needs to register two DNS hostnames: one DNS hostname for the [MinIO Console](#), and another one for the [MinIO API endpoint](#).

IM will deploy a Virtual Machine with Docker and MinIO as a containerised service. Therefore, the first step for the user is to select [Deploy a VM](#) in IM:





UCA TESTBED INFRASTRUCTURE

Examind / Iceberg/WebODV effective from this month end.

More precise elements:

Cloud IaaS : Openstack (community version)

<https://www.openstack.org/>

total : 48 hypervisors, 800 CPU oores, RAM 7.6 TB,

Distributed storage/ S3 : Ceph (community version) <https://ceph.io/>

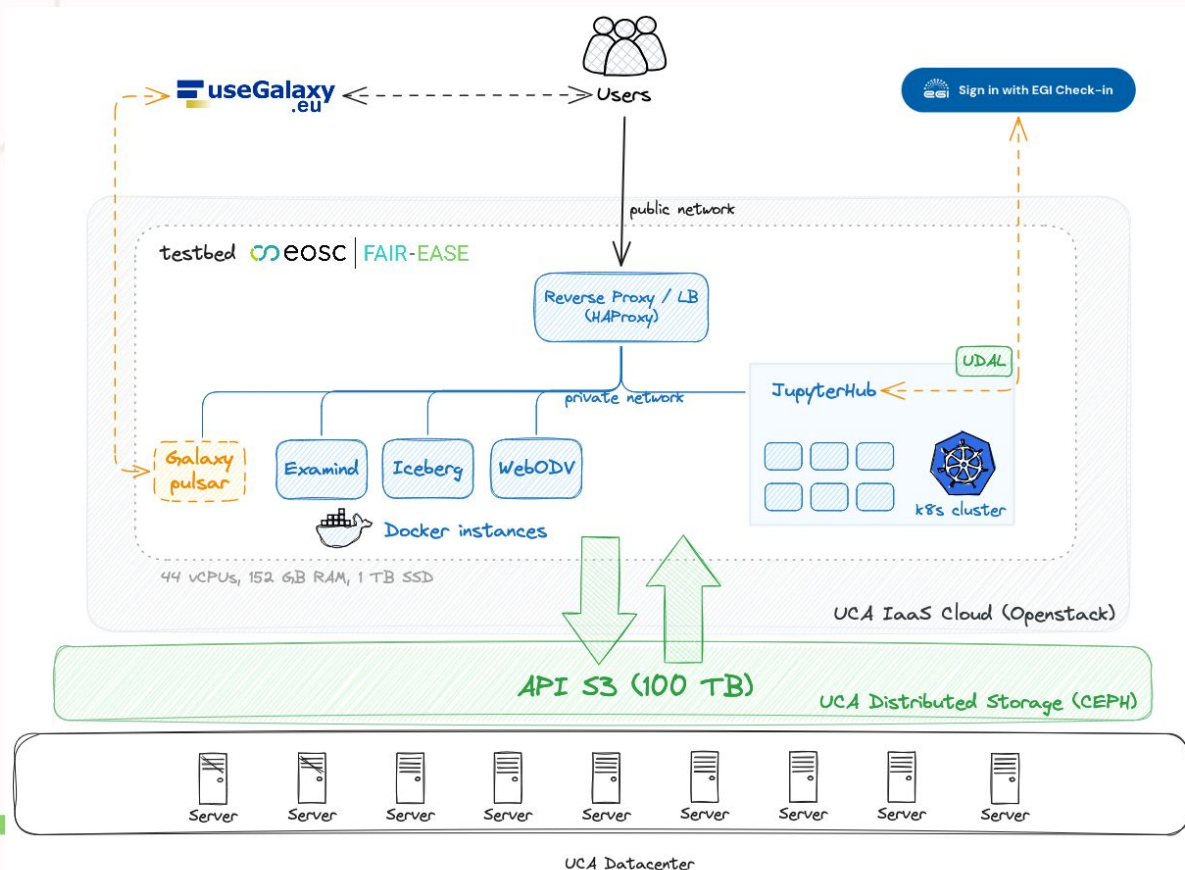
total : 600 TB SSD (for vm openstack), 5.5 PB HDD (S3)
(replicated 3 times so /3)

FAIR-EASE specifically allocated resources: 44 CPUs, 152 GB RAM, 128 GB SSD, 100 TB via API S3

Cluster kubernetes FE : kubernetes v1.31, déployed with Terraform et Talos Linux (<https://www.talos.dev/>)

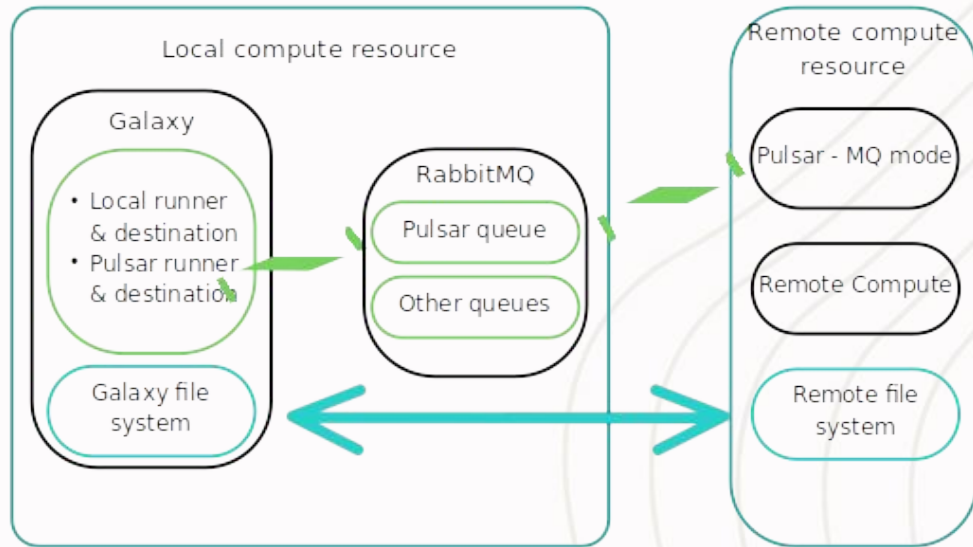
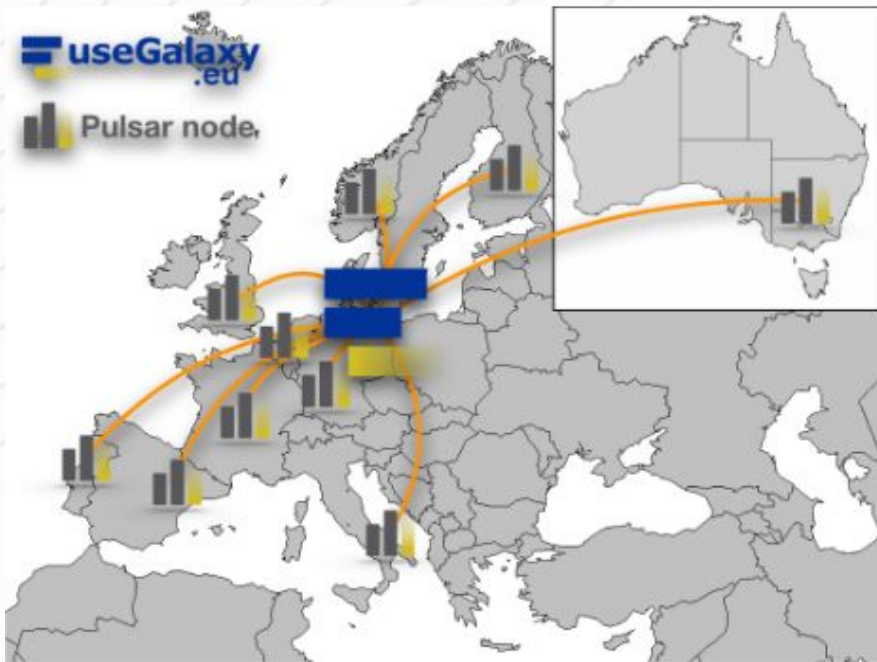
3 vm for control plane

3 vm for workers executed on client services instances (JupyterHub/JupyterLab)





Galaxy deals with remote compute resources and remote files system



↔ AMQP messages

↔ File transport (curl)

Apache Iceberg an open table format for huge analytic datasets (including metadata)

Iceberg adds tables to compute engines including Spark, Trino, PrestoDB, Flink, Hive and Impala using a high-performance table format that works just like a SQL table. <https://iceberg.apache.org/docs/1.5.2/>

- Iceberg avoids unpleasant surprises. Schema evolution works and won't inadvertently un-delete data. Users **don't need to know about partitioning to get fast queries.**
- Iceberg is used in production where a **single table can contain tens of petabytes** of data and even these huge tables can be read without a distributed SQL engine.
- Open standard designed and developed to be an open community standard with a specification to ensure compatibility across languages and implementations. **Apache Iceberg is open source**, and is developed at the Apache Software Foundation.

Summary

Iceberg excels in decoupling storage from compute, managing evolving schemas, and enabling scalable, FAIR-compliant data workflows, making it ideal for research infrastructures.

For FAIR, open data infrastructures like FAIR-EASE, Iceberg offers flexibility and adaptability while ensuring long-term usability and performance at scale.

Iceberg to ERDDAP data publication

UCA S3 Object Store

<https://s3.mesocentre.uca.fr>



[s3://eosc-fairease-warehouse](https://s3.mesocentre.uca.fr)



Catalog

+ pySpark
(Query Engine)



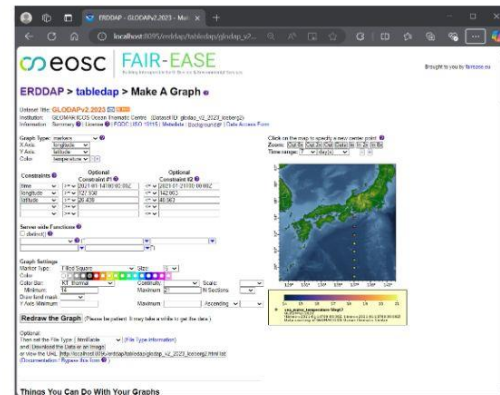
Trino Clients SQL interfaces



Trino (Query Engine)

JDBC Driver

ERDDAP.eosc-fairease



Integrating New Data Storage and Access Paradigms in pilot's practices

The top screenshot shows the 'Object Browser' interface for a 'warehouse' bucket. It displays a list of objects with columns for Name, Last Modified, and Size. The bottom screenshot shows a JupyterLab notebook with SQL queries and their results.

Object Browser Data:

Name	Last Modified	Size
00003-4-2f596d6-0968-487c-8390-5cc32b645635-0-00001.parquet	Today, 10:12	51.5 MiB
00009-10-2f596d6-0968-487c-8390-5cc32b645635-0-00001.parquet	Today, 10:12	25.9 MiB

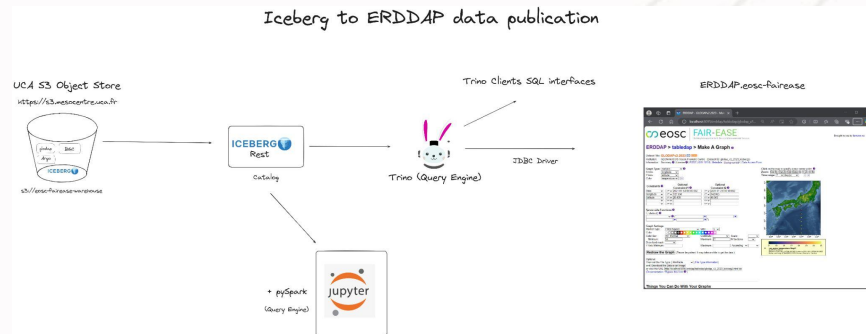
JupyterLab SQL Query Results:

```
SELECT COUNT(*) as cnt
FROM glodap.glodap_v2_2023
```

cnt: 1402829

```
SELECT doi,latitude,longitude,depth,MAX(temperature) FROM glodap.glodap_v2_2023 GROUP BY doi,latitude,longitude,depth ORDER BY MAX(temperature) DESC;
```

doi	latitude	longitude	depth	max(temperature)
https://doi.org/10.25921/16y6-9e29	-18.0	-71.92	6.0	34.501
https://doi.org/10.25921/16y6-9e29	-22.0	-70.75	17.0	32.975
https://doi.org/10.25921/0sta-y820	16.0	41.5833	0.0	32.68
https://doi.org/10.25921/0sta-y820	16.0	41.5833	10.0	32.63
https://doi.org/10.25921/0sta-y820	16.0	41.5833	20.0	32.5
https://doi.org/10.25921/0sta-y820	14.7	42.1667	0.0	32.4
https://doi.org/10.25921/0sta-y820	14.7	42.1667	10.0	32.39
https://doi.org/10.25921/0sta-y820	14.7	42.1667	20.0	32.35
https://doi.org/10.3334/cdiac/otg.ndp080	24.569	57.222	4.0	32.171
https://doi.org/10.3334/cdiac/otg.ndp080	26.411	56.57	4.0	31.912



"Open Data Lake House Format - Apache Iceberg - Quick start" Damian Smyth:

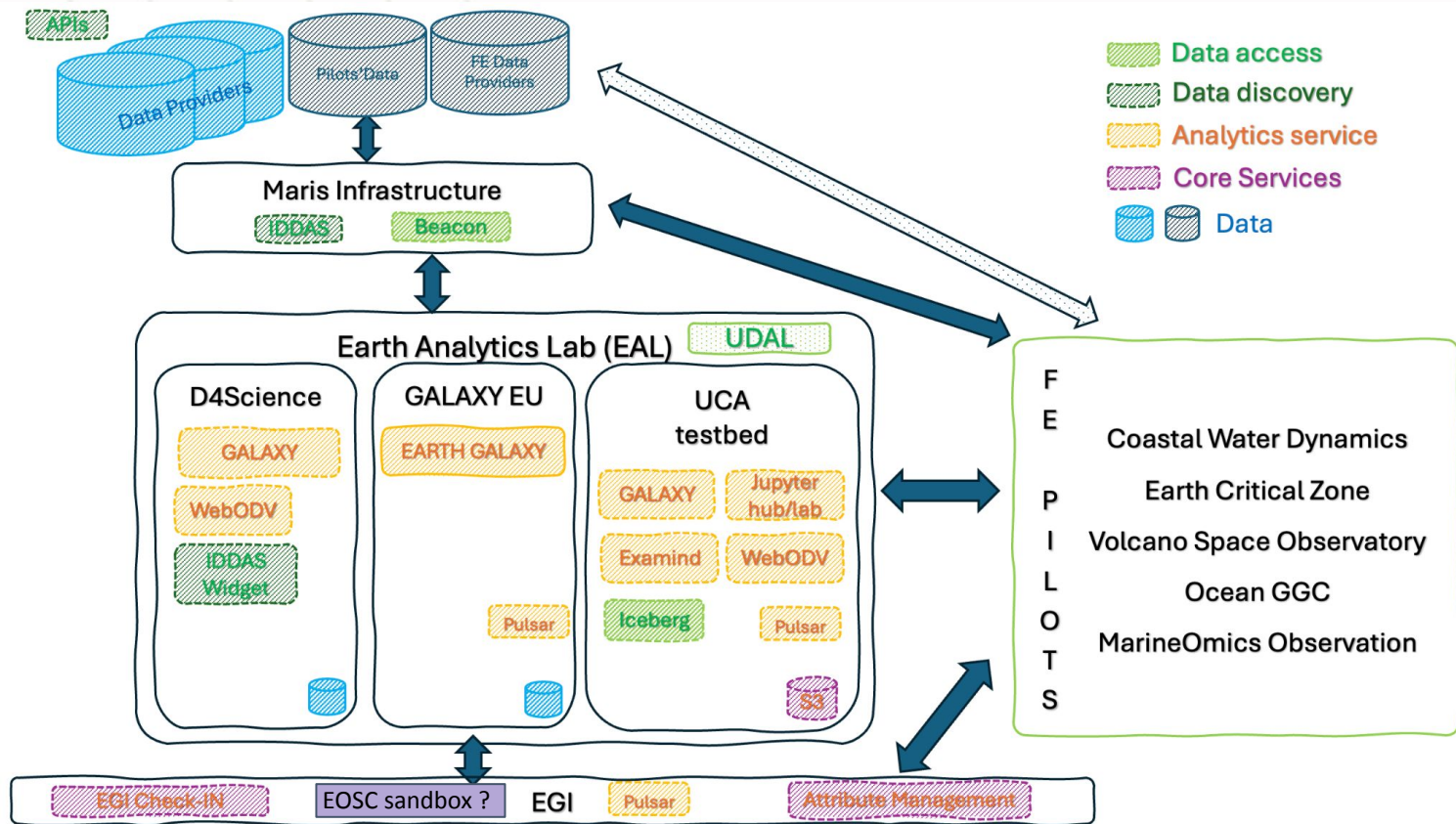
The Examind provides all the features you need to build a Geographic, Hydrographic, Oceanographic and Meteorologic geospatial data processing infrastructure. Designed for **interoperability**, all the products in the suite comply with **OGC standards**, and integrate OGC Web services

Examind community is a **open-source platform / server**.

- It manages multiple data formats (Netcdf, geotiff, etc.), clouds native data formats (COGs, GIMI), multiple OGC (Open Geospatial Consortium) standards (WCS, WMS, WPS, OGC API, etc.)
- It has a number of ready-to-use processes for various uses in the geospatial world.
- It offers several ways of managing and structuring data. The server can connect to an existing data source, hold the data locally, or generate new data via different processes (via WPS, OGC API Process, openEO); all via different paradigms, such as data cube structuring.

Examind supports several data storage options local storage, FTP, S3 (AWS / Minio), via HTTP/HTTPS, or storage from a WMS / WMTS service, etc.

The service can be deployed via a docker image and a docker-compose file on kubernetes infrastructures.



Summary

Iceberg excels in decoupling storage from compute, managing evolving schemas, and enabling scalable, FAIR-compliant data workflows, making it ideal for research infrastructures.

For FAIR, open data infrastructures like FAIR-EASE, Iceberg offers flexibility and adaptability while ensuring long-term usability and performance at scale.



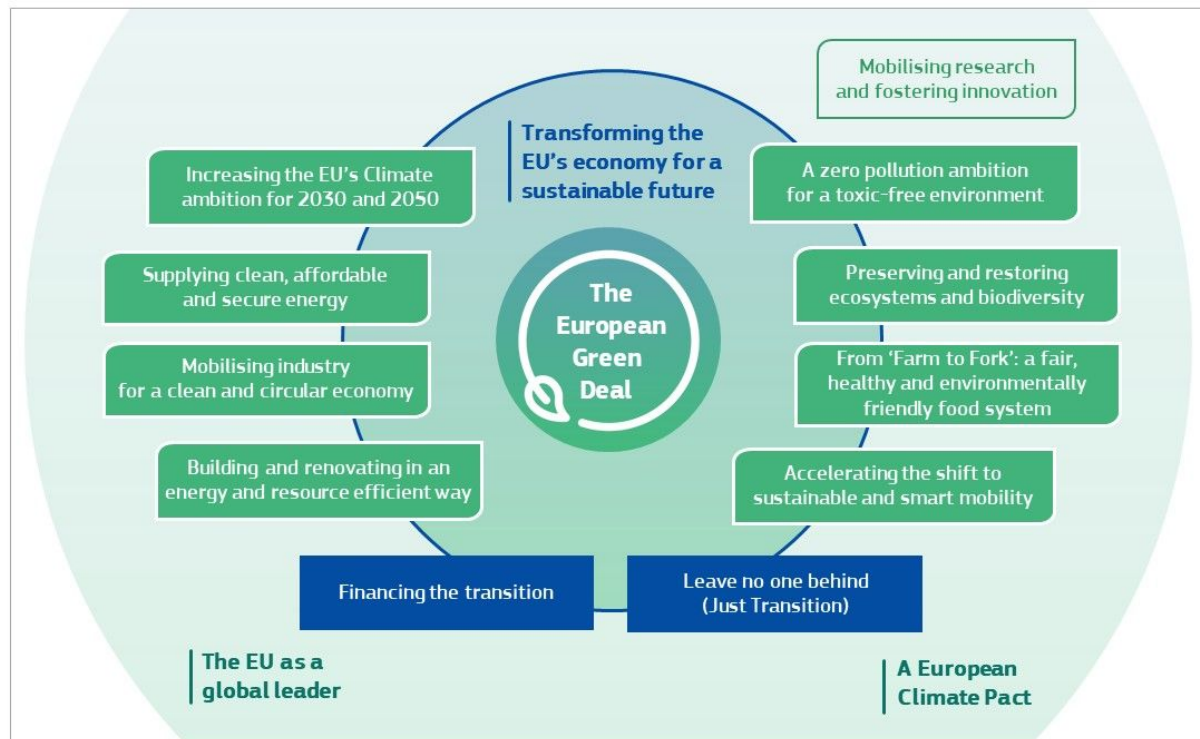


WEDNESDAY AFTERNOON



FAIR-EASE in the context of the European green deal





FAIREASE in the context of the european green deal

European green deal topics

- **Designing a set of deeply transformative policies**
- **Mainstreaming sustainability in all EU policie**

- 2.1.7.
- *Preserving and restoring ecosystems and biodiversity*
- 2.2.2.
- *Greening national budgets and sending the right price signals*
- 2.2.3.
- *Mobilising research and fostering innovation*

- 2.1.7.
- *Preserving and restoring ecosystems and biodiversity*
- 2.2.2.
- *Greening national budgets and sending the right price signals*
- 2.2.3.
- *Mobilising research and fostering innovation*

<https://www.aeris-data.fr/leruption-volcanique-sur-les-iles-tonga-vue-par-les-satellites-geostationnaires/>

- 2.1.7.
- Preserving and restoring ecosystems and biodiversity
- 2.2.2.
- Greening national budgets and sending the right price signals
- 2.2.3.
- Mobilising research and fostering innovation

<https://www.aeris-data.fr/le-vue-par-les-satellites-geostationnaires/>

