27-29 May 2024

International conference on Marine Data and Information Systems

SeaDataNet AISBL · Havforskningsinstituttet · Ifremer · MARIS · MEDIN · National Oceanography Centre · INGV · iMagine · IODE · eosc Blue-Cloud2026

eosc | FAIR-EASE

The **GALAXY** platform as a ground breaking **FAIR** tool for **EARTH SYSTEM**'s analytical **WORKFLOWS**

Marie Jossé

Jérôme Detoc

Erwan Bodéré

Funded by the European Union

OVERVIEW

01

Fair-Ease

02

The Earth System

03

Why Galaxy ?

04

Good pratices

05

Our tools and workflows

06

Try it !

eosc | FAIR-EASE

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

THE EARTH SYSTEM

By Fair-Ease

The Earth System is a complex and dynamic system that encompasses the interactions between the atmosphere, oceans, land, and biosphere.

Five pilots concerning applications for real-life science use-cases came with their specific needs in virtual research environments

Coastal Water Dynamics

Focuses on the coastal marine environment near river estuaries, where important processes take place.

Earth Critical Zone

Improve analytics model specific to the earth critival zone. Monitors land and soil degradation.

Volcano space observatory

Monitors global volcanic activity, allowing the focus on any major volcanic eruption worldwide. Aggregating remote sensing observations from the solid earth and the atmosphere

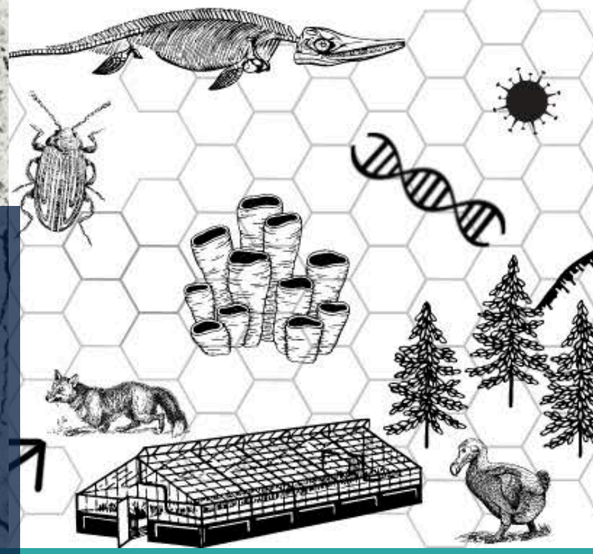
Ocean Bio-Geochemical

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

Marine Omics

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



WHY GALAXY

A community driven platform popular in other domains

Reusing the existing : 2 subdomains of Galaxy europe inspiring Fair-Ease, the climate and ecology ones

---

FAIR by design

Opensource web platform for sharing and processing research data

- Accessibility
- Reproducibility
- Transparency
- Community work

Tools

search tools

Upload Data

- Get Data
- Send Data
- Collection Operations

GENERAL TEXT TOOLS

- Text Manipulation
- Filter and Sort
- Join, Subtract and Group
- Convert Formats

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

WORKFLOWS

- All workflows

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

History

Rechercher des données

Ocean's variables

#training

206 MB | 21

- 35: Interactive JupyTool and not ebook on data 32
- 34: Executed JupyTool Notebook
- 33: JupyTool output collection
  a list with 0 datasets
- 32: CMEMS OPeNDAP TEST.ipynb
- 31: CMEMS MOTU Test.ipynb
- 30: CMEMS FTP Test.ipynb
- ⚠ 29: Executed Holoviz
- ⚠ 28: Holoviz notebook output collection
  a list with datasets
- 14: openEO_2022-05-28Z
- 13: openEO_2022-05-18Z
- 12: openEO_2022-05-15Z

Galaxy

eosc | FAIR-EASE

GALAXY FEATURES

FAIR

Galaxy

Tool name, id, description, ...

Open source platform for everyone

One tool for multiple kind of datasets

To be chained in different workflows

Best practices

Atomisation

Divided in elementary bricks

Generalisation

Standardised tools

Riguous

Peer review

```xml
<tool id="obis_data" name="OBIS occurences" version="@VERSION@" profile="20.01">
    <description>retrieve data</description>
    <macros>
        <import>macro_obis.xml</import>
    </macros>
    <expand macro="topic"/>
    <expand macro="requirements">
        <requirement type="package" version="2.11.3">r-robis</requirement>
    </expand>
    <required_files>
        <include type="literal" path="robis.r"/>
    </required_files>
```

```xml
    <command detect_errors="exit_code"><![CDATA[
        Rscript
            '$__tool_directory__/robis.r'
            '$species'
            '$taxon'
            '$lat_min'
            '$lat_max'
            '$long_min'
            '$long_max'
            '$output'
        ]]>
```

```xml
    <inputs>
        <param name="species" type="text" format="character" label="Scientific name of the sp
            <validator type="regex">^[A-Za-z ]*$</validator>
        </param>
        <param name="taxon" type="text" label="Taxon ID">
            <validator type="regex">^[0-9]*$</validator>
        </param>
        <param name="lat_min" type="float" min="-90" max="90" value="0" label="Input latitude
        <param name="lat_max" type="float" min="-90" max="90" value="0" label="Input latitude
        <param name="long_min" type="float" min="-90" max="90" value="0" label="Input longitu
        <param name="long_max" type="float" min="-90" max="90" value="0" label="Input longitu
    </inputs>
```

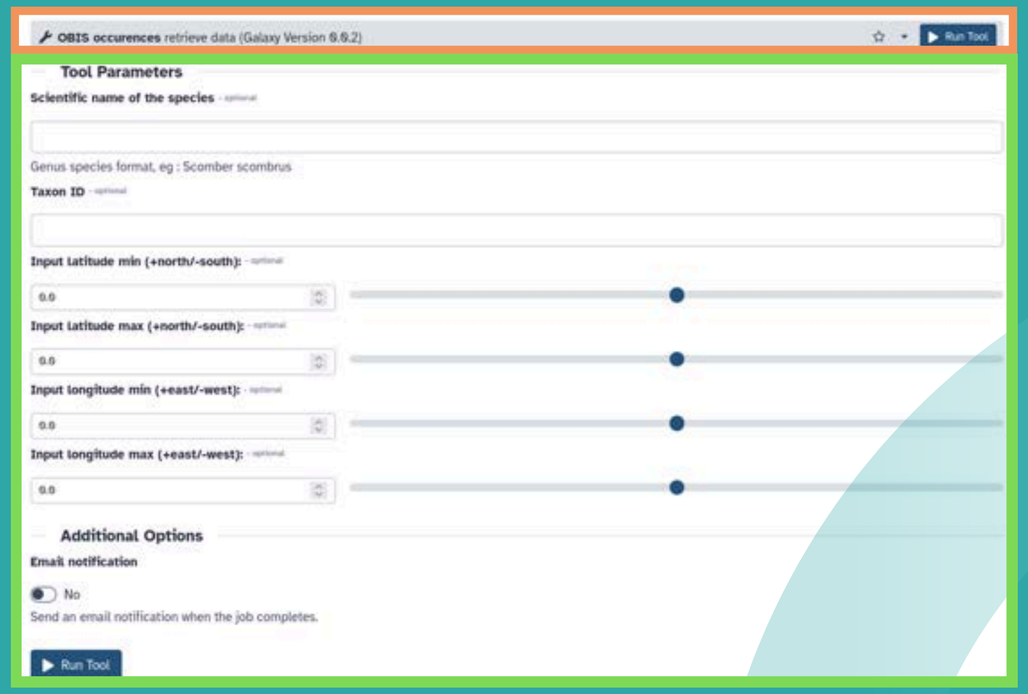
```xml
    <outputs>
        <data name="output" format="tabular" from_work_dir="output.tab" label="Species occure
    </outputs>
```

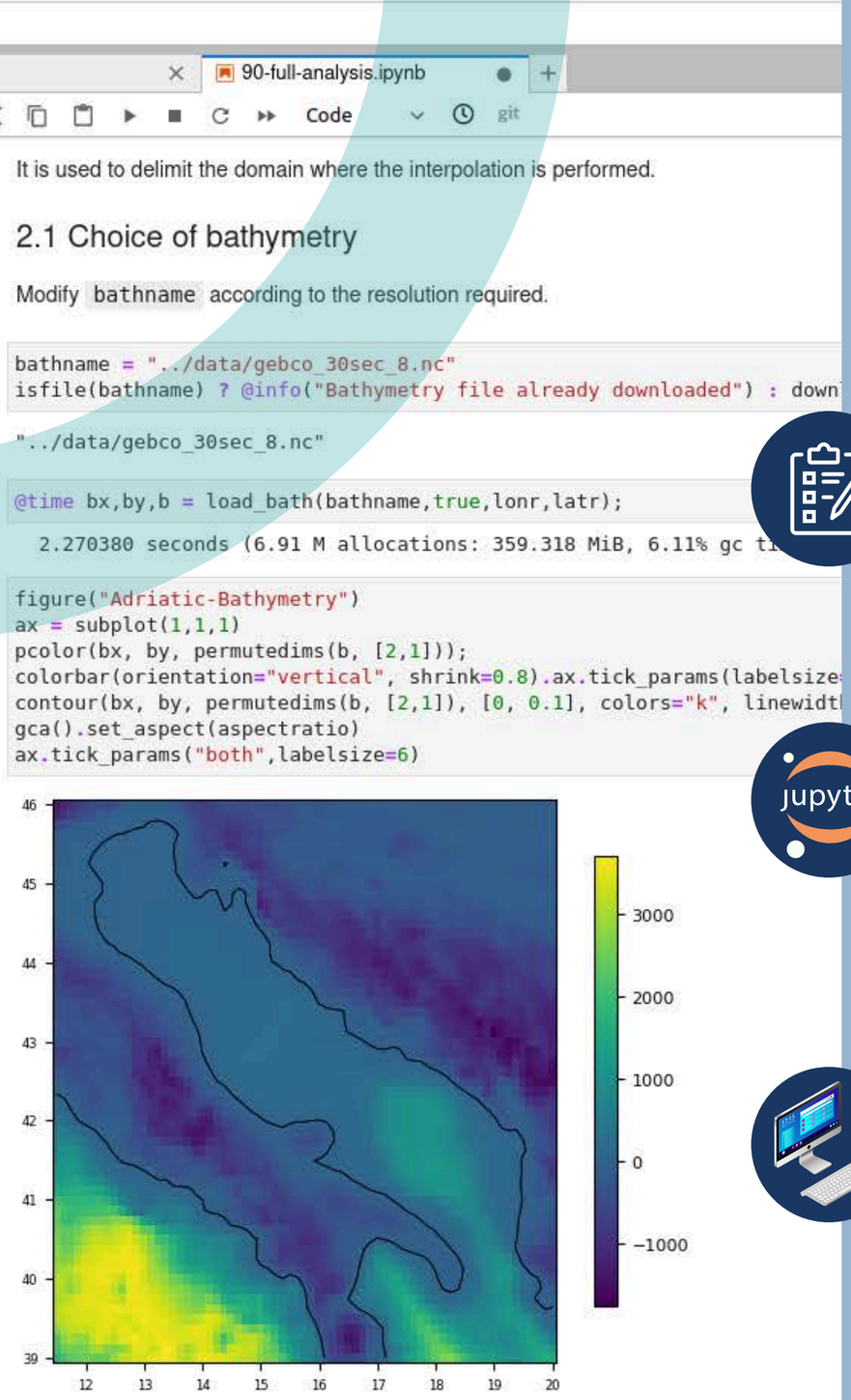
```xml
    <tests>
        <test expect_num_outputs="1">
            <param name="species" value="Scomber scombrus"/>
            <output name="output">
                <assert_contents>
                    <has_text text="Scombridae"/>
                    <has_n_columns n="163"/>
                </assert_contents>
            </output>
        </test>
        <test expect_num_outputs="1">
            <param name="lat_min" value="6"/>
```

Standardisation

AN XML FILE FOR EACH TOOL

- Each script (python, R, java …) is linked to a “wrapper”, the xml file, where some necessary info must be written.





GALAXY & TOOLS

Batch tools

Inputs parameters into a formular, execute the tool and directly retrieve the wanted outputs. **NO need for any informatic skills**

Jupyterlab tools

Interactive tools that can be launched with specific environment and set of notebooks (DIVAnd, Holoviz, Copernicus Data Space Ecosystem, Pangeo, ...)

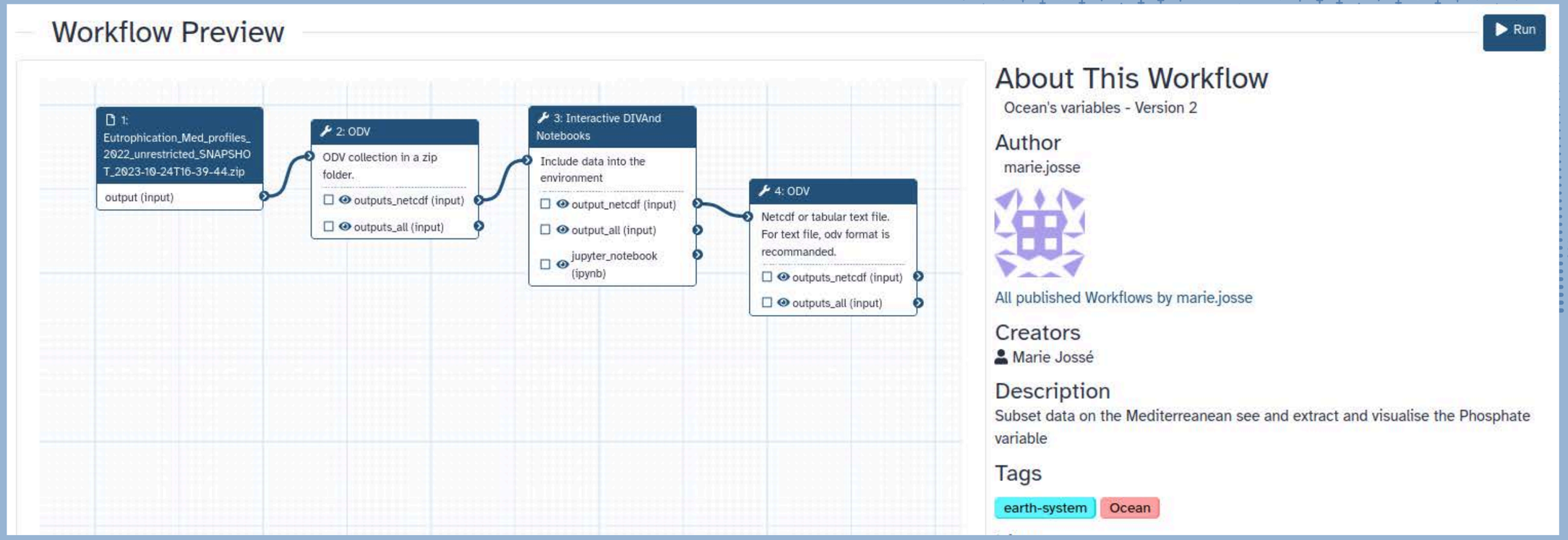
Desktop tools

Intercative tools that can be launched through Galaxy and allow you to have desktop applications launched in a browser window (QGIS, ODV, ...)

Galaxy

Atomisation

CREATE WORKFLOWS FROM ELEMENTARY BRICKS



Allows better understanding of an analysis

Smaller scripts easier to read

Allows researchers to combine tools

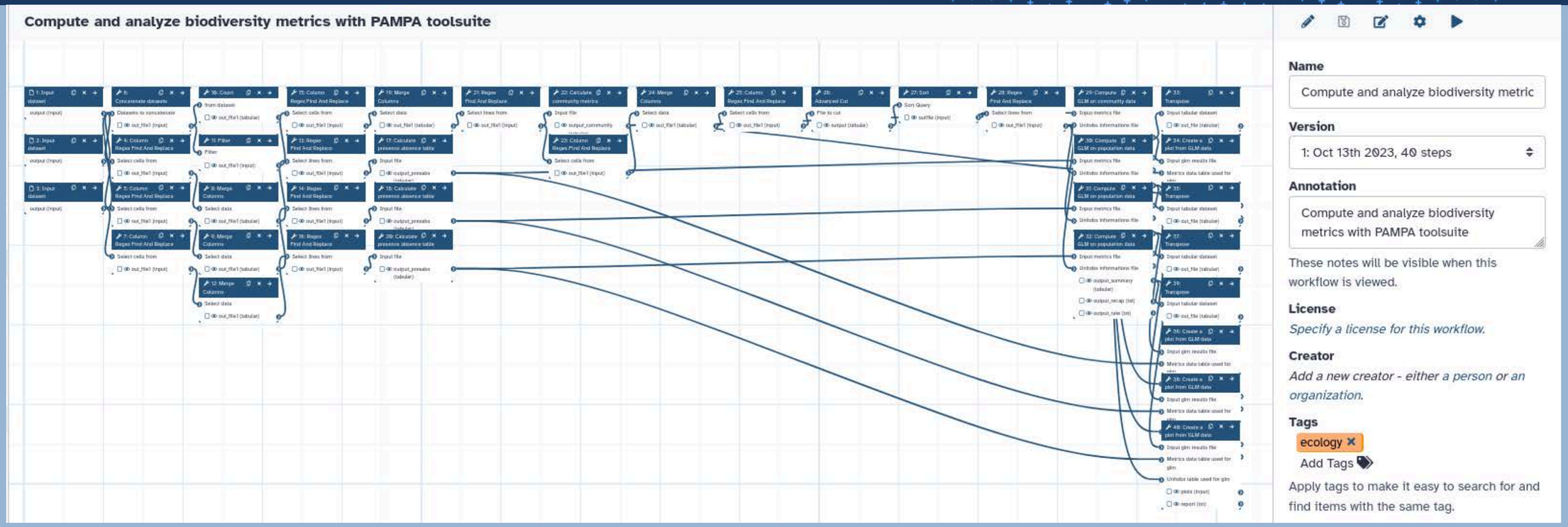
Can combine tools to have a full operationnal workflow

Allows to create different kind of analysis

Can use the tool with different ones and have multiple workflows

Atomisation

CREATE WORKFLOWS FROM ELEMENTARY BRICKS



Allows better understanding of an analysis

Smaller scripts easier to read

Allows researchers to combine tools

Can combine tools to have a full operationnal workflow

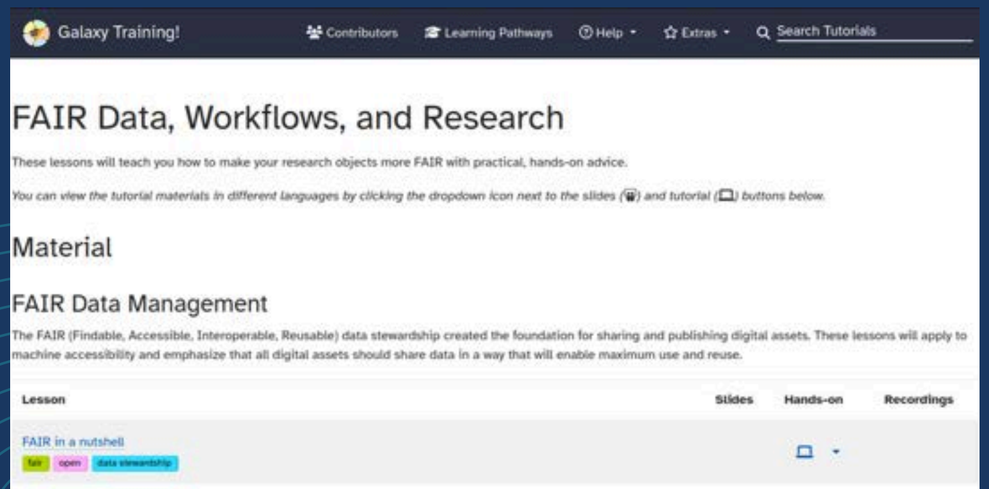
Allows to create different kind of analysis

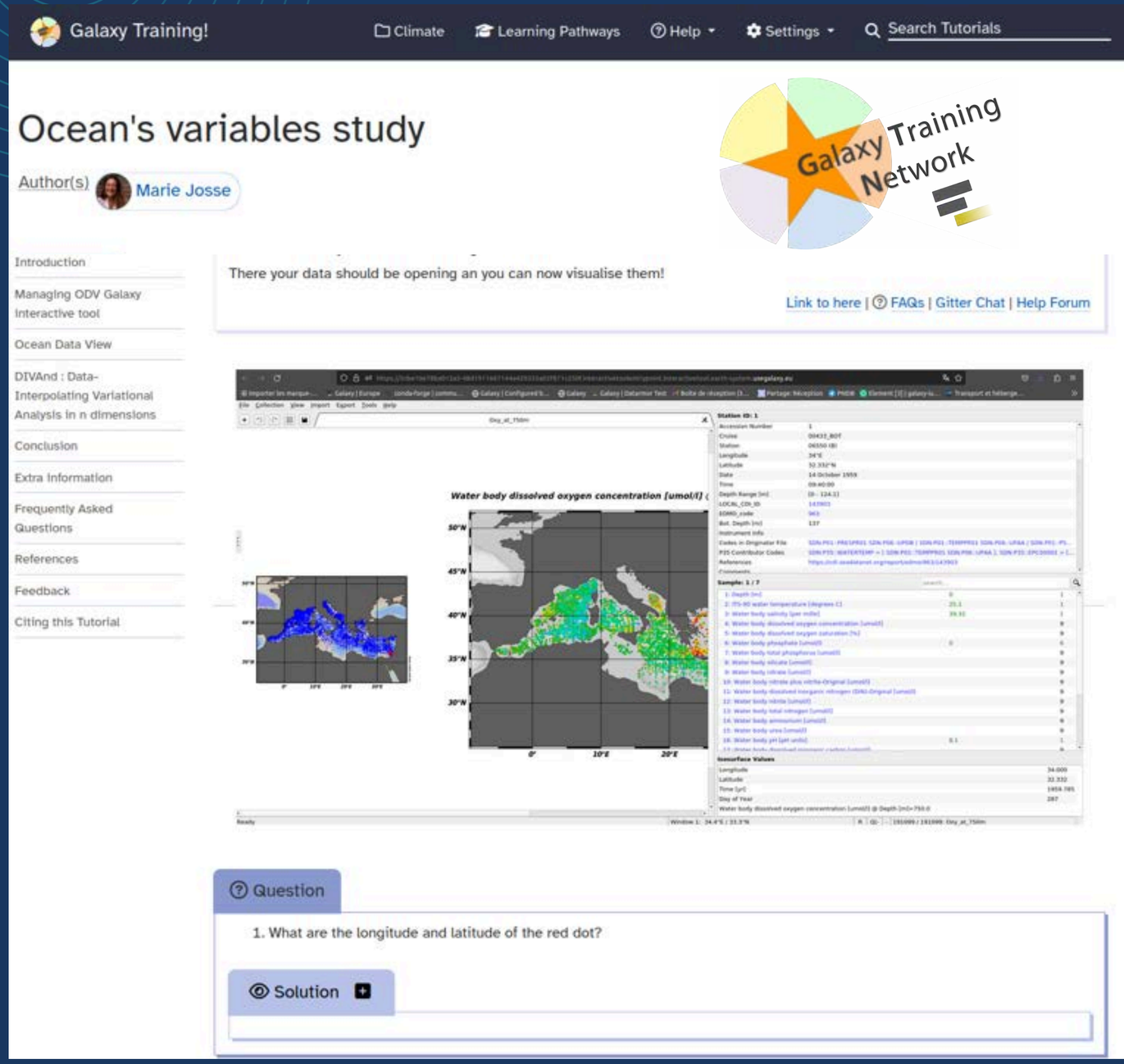
Can use the tool with different ones and have multiple workflows

The Galaxy Training Network

- Tutorials
- Classes and courses
- Pathways

- An easy way to learn galaxy and improve your skills on various domains for instance a set of tuto are available on FAIR management





GALAXY JOURNEY STEPS

Create tools

As explained you can easily create tools (interactive or not) with an xml file

Set up a workflow

- Can be shared, published, runned, ...
- Workflows with human in the loop : jupyterlab, desktop applications, ...
- Fully automatic workflows : click button tools without any knowledge in programmation

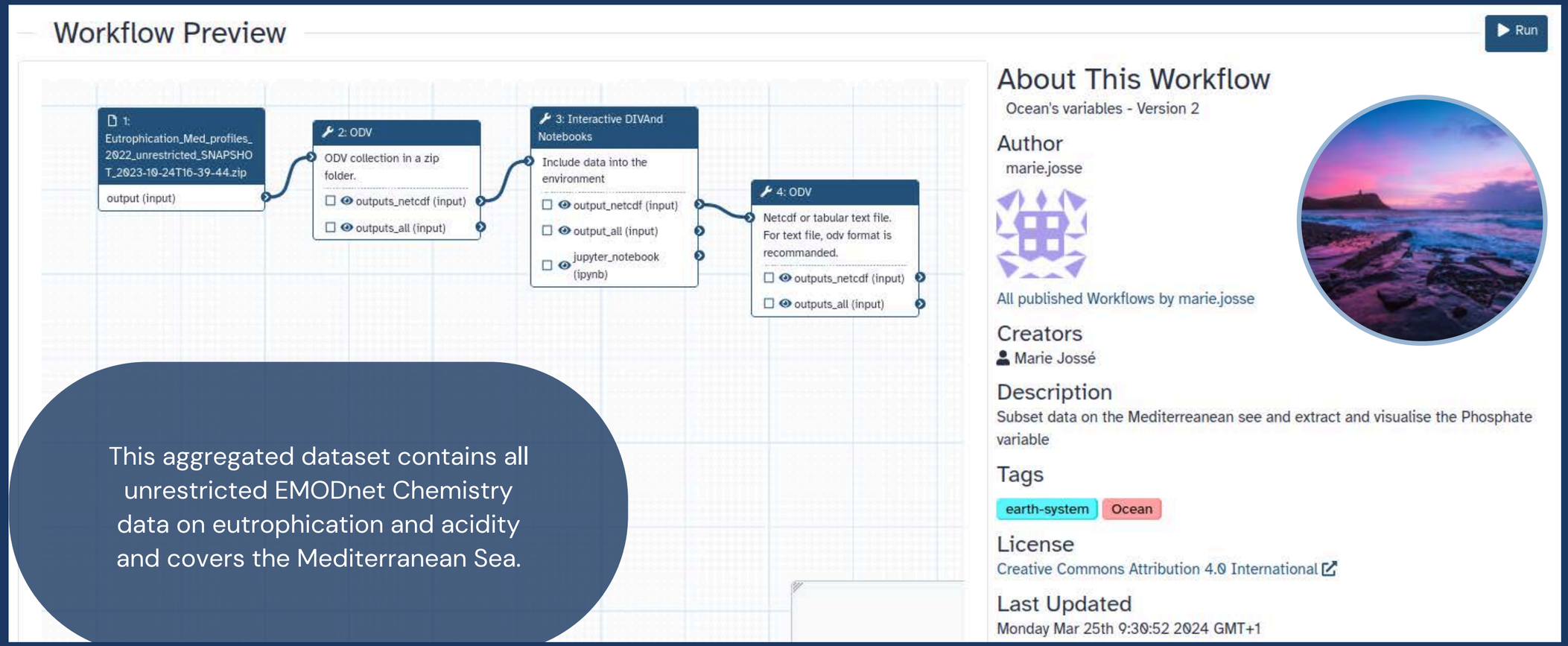
Write a tutorial

- Easily wrote in markdown
- Lot of guides on how to write the perfect tutorial
- Can even be extracted from the workflow.

eosc | FAIR-EASE

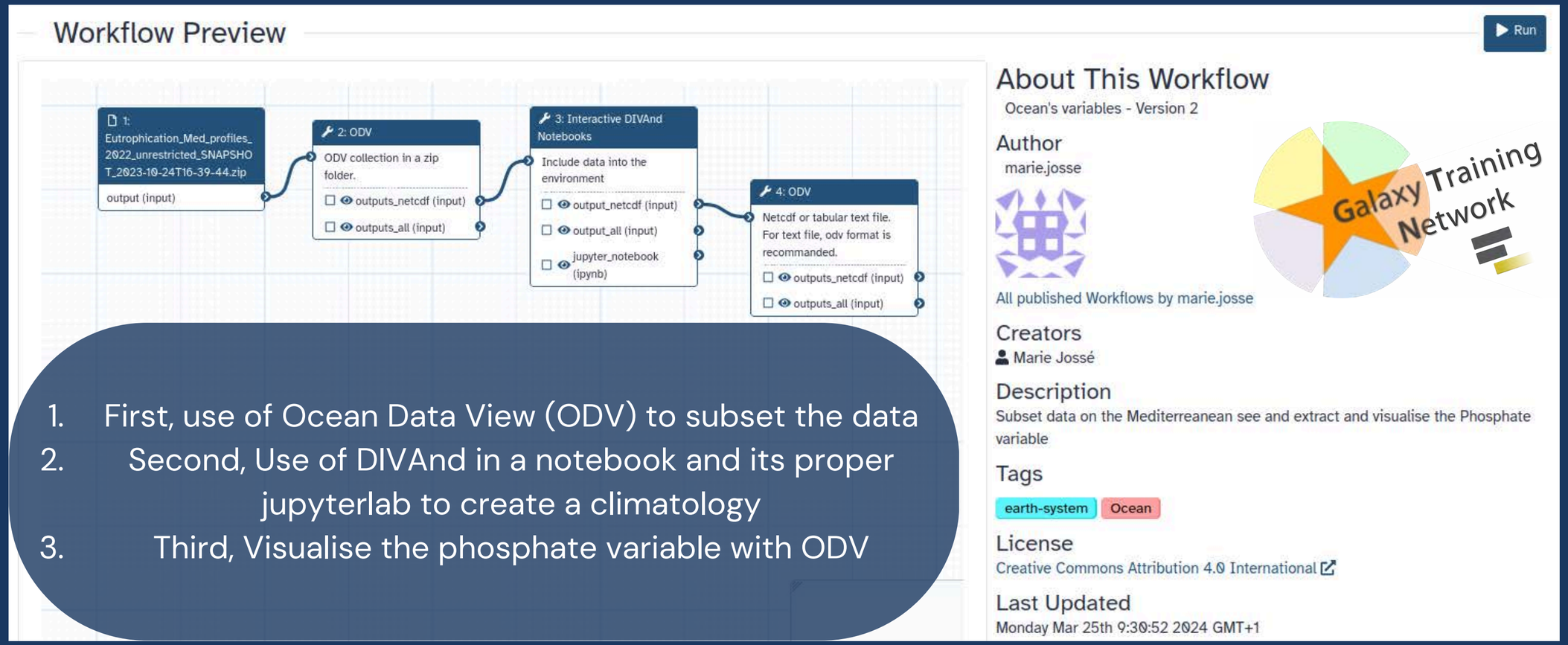
OUR TOOLS AND WORKFLOWS

OCEAN VARIABLES STUDY :
Ocean Data View & DIVAnd



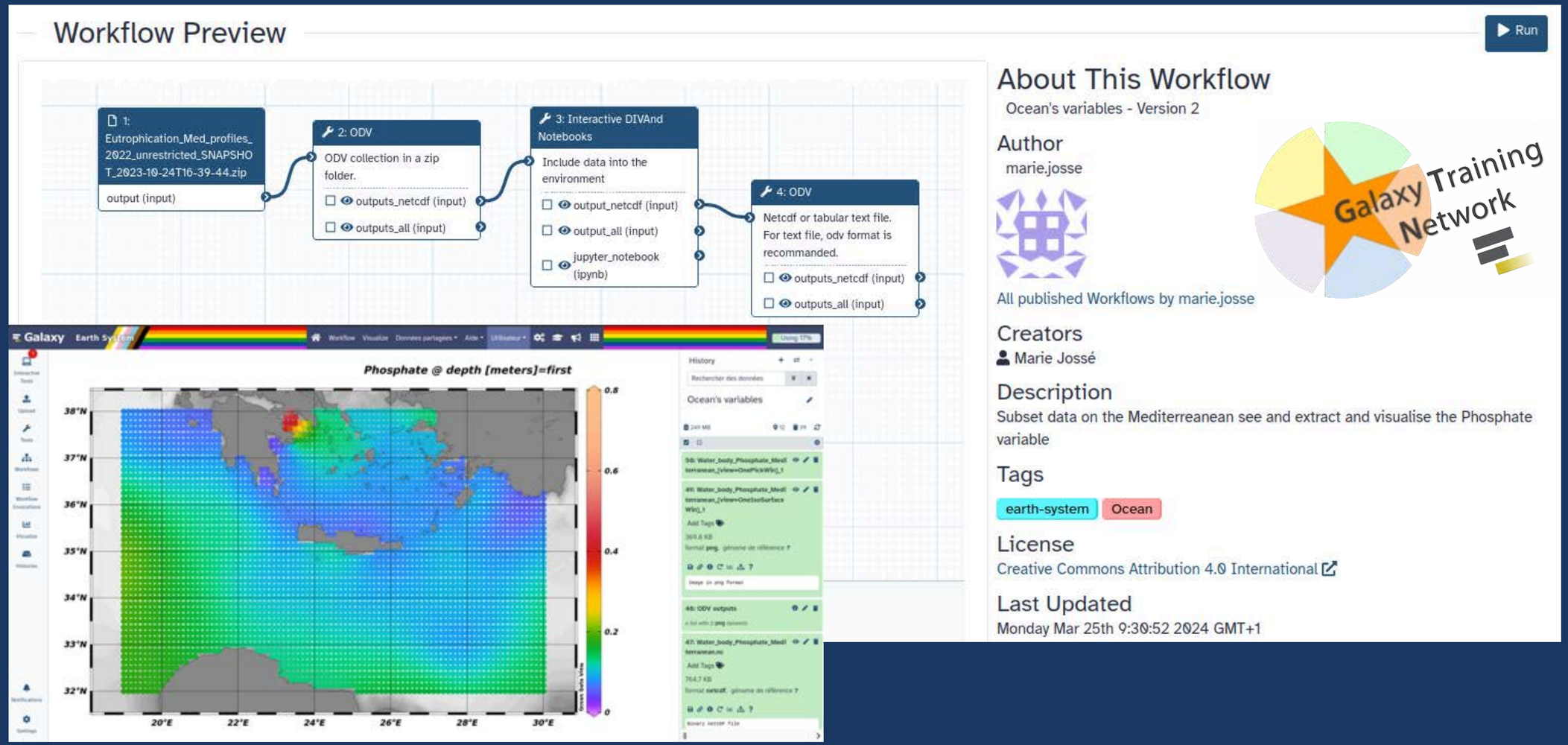
This aggregated dataset contains all unrestricted EMODnet Chemistry data on eutrophication and acidity and covers the Mediterranean Sea.

OCEAN VARIABLES STUDY :
Ocean Data View & DIVAnd



1. First, use of Ocean Data View (ODV) to subset the data
2. Second, Use of DIVAnd in a notebook and its proper jupyterlab to create a climatology
3. Third, Visualise the phosphate variable with ODV

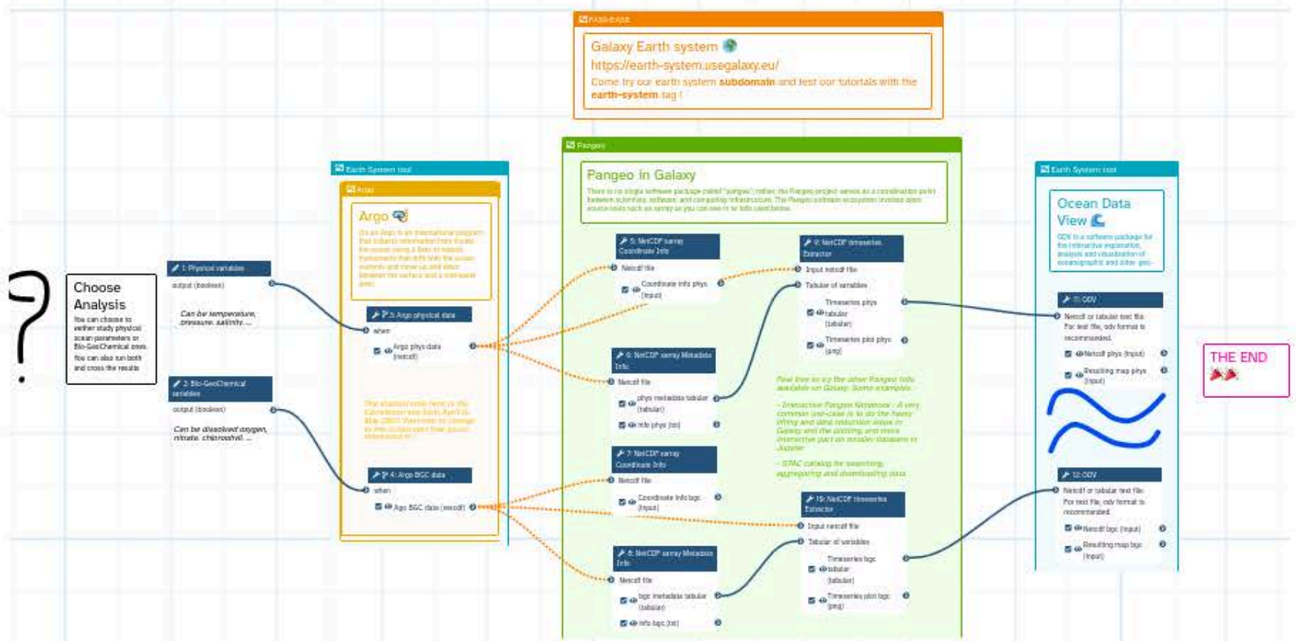
OCEAN VARIABLES STUDY :
Ocean Data View & DIVAnd

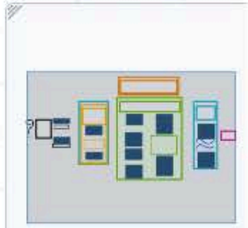


ANALYSE ARGO DATA WITH PANGEO TOOLS & ODV

Workflow Preview

Run





About This Workflow

Full Analyse Argo data - Version 14

Author

marie.josse

All published Workflows by marie.josse

Creators

Marie Jossé

Description

Process argo data with the Pangeo Ecosystem and visualise them with Ocean Data View (ODV)

Tags

Ocean | Pangeo | earth-system

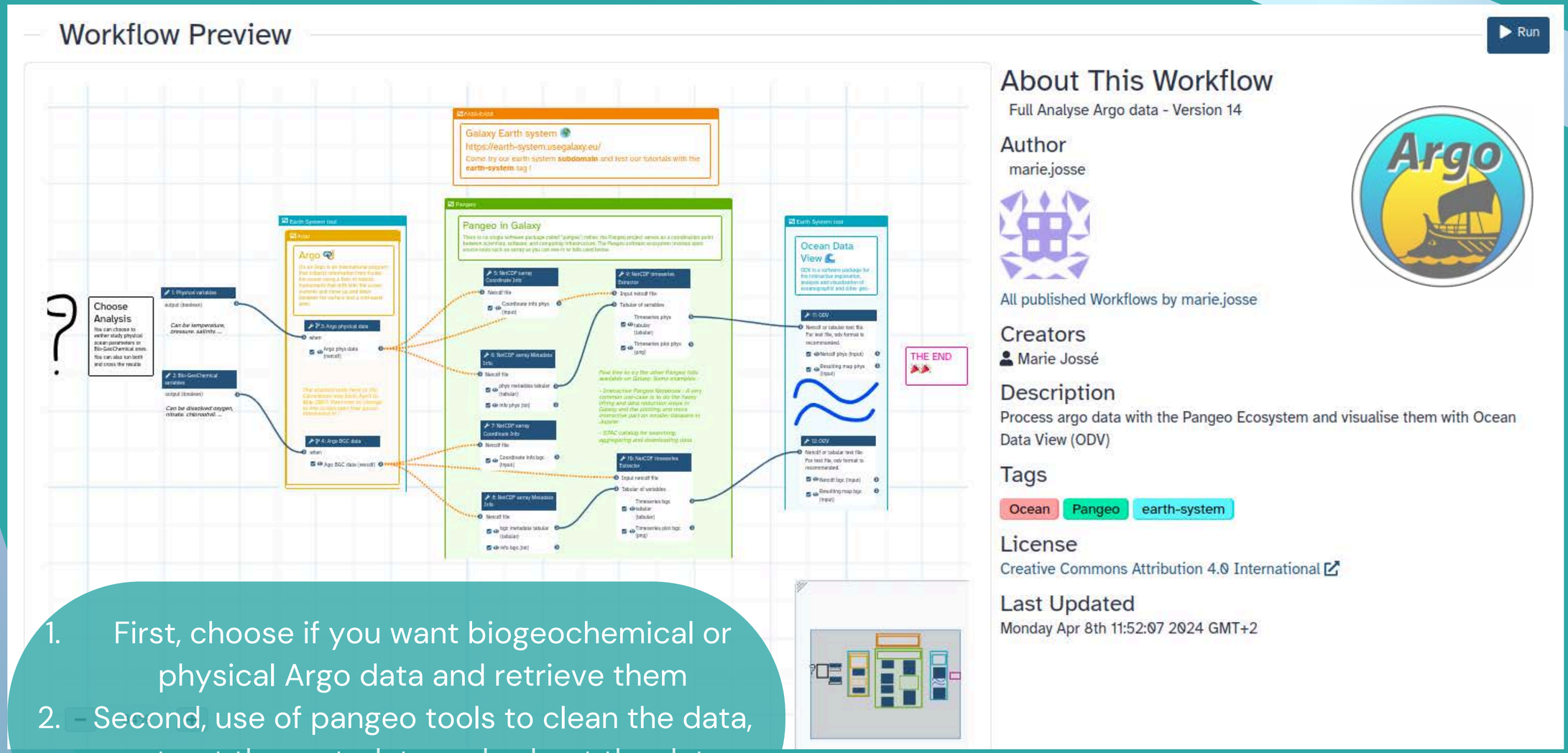
License

Creative Commons Attribution 4.0 International

Last Updated

Monday Apr 8th 11:52:07 2024 GMT+2

ANALYSE ARGO DATA WITH PANGEO TOOLS & ODV

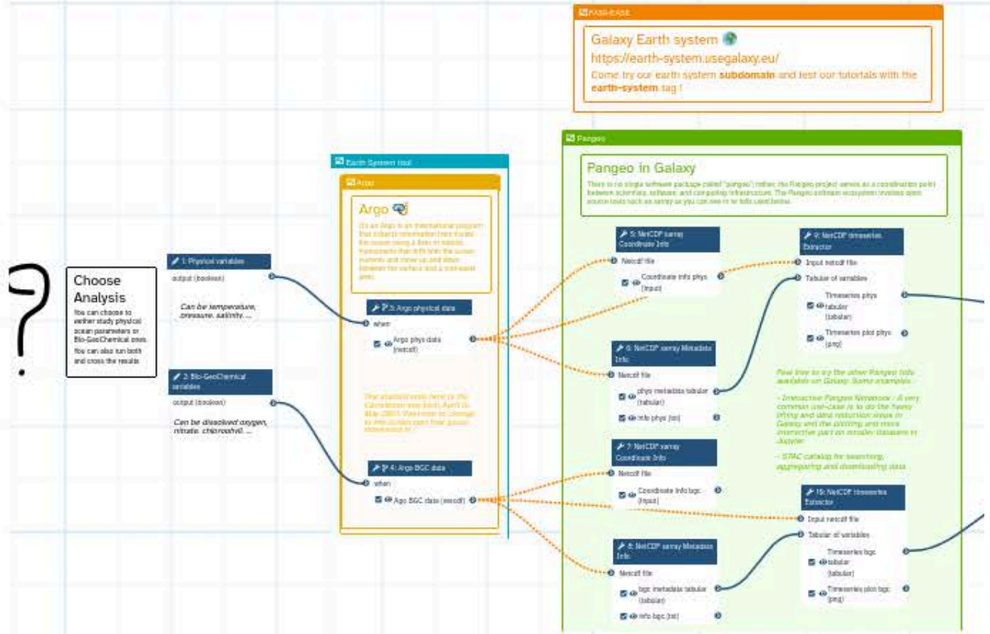


1. First, choose if you want biogeochemical or physical Argo data and retrieve them
2. Second, use of pangeo tools to clean the data, extract the metadata and subset the data
3. Third, Visualise the Argo data with ODV

ANALYSE ARGO DATA WITH PANGEO TOOLS & ODV

Workflow Preview

Run



Argo data analysed, managed, and visualised by the Pangeo tools

Visualisation mapping with a GaLaxy Earth System's tool

Ocean Data View (ODV)

Conclusion

Extra Information

Frequently Asked Questions

Feedback

Citing this Tutorial

About This Workflow

Full Analyse Argo data - Version 14

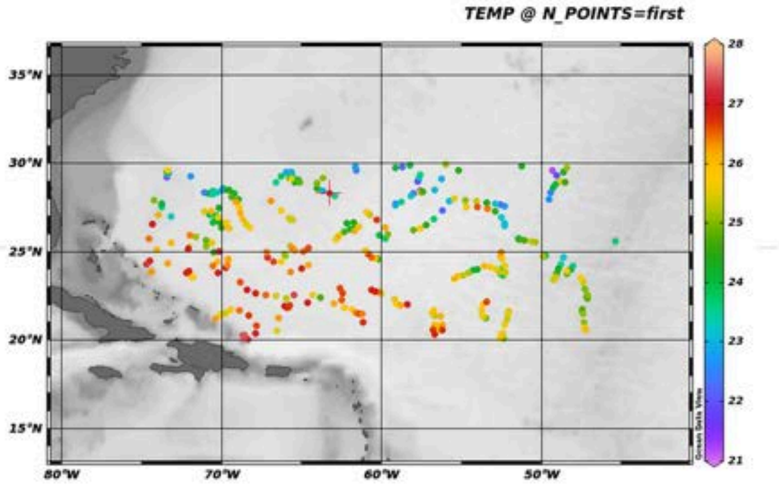
Author

marie.josse

Hands-on: Save your visualisation map

1. Click right on the map select **Save Plot As...**
2. In the pop-up screen go to the folder **ODV**, **galaxy**, **outputs**.
3. In **File name** rename your view (for example subset_temp_argo_data)
4. In **Files of type** select `PNG (*.png *.PNG)` and **Save** then **OK** and **OK**. Once you're finished with ODV :
5. On th top left click on **File** select **Exit**
6. If you want to save the other window also click on **Yes**. Here we don't need it so click **No**.

Link to here | FAQs | Gitter Chat | Help Forum



What you need to keep in mind !

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

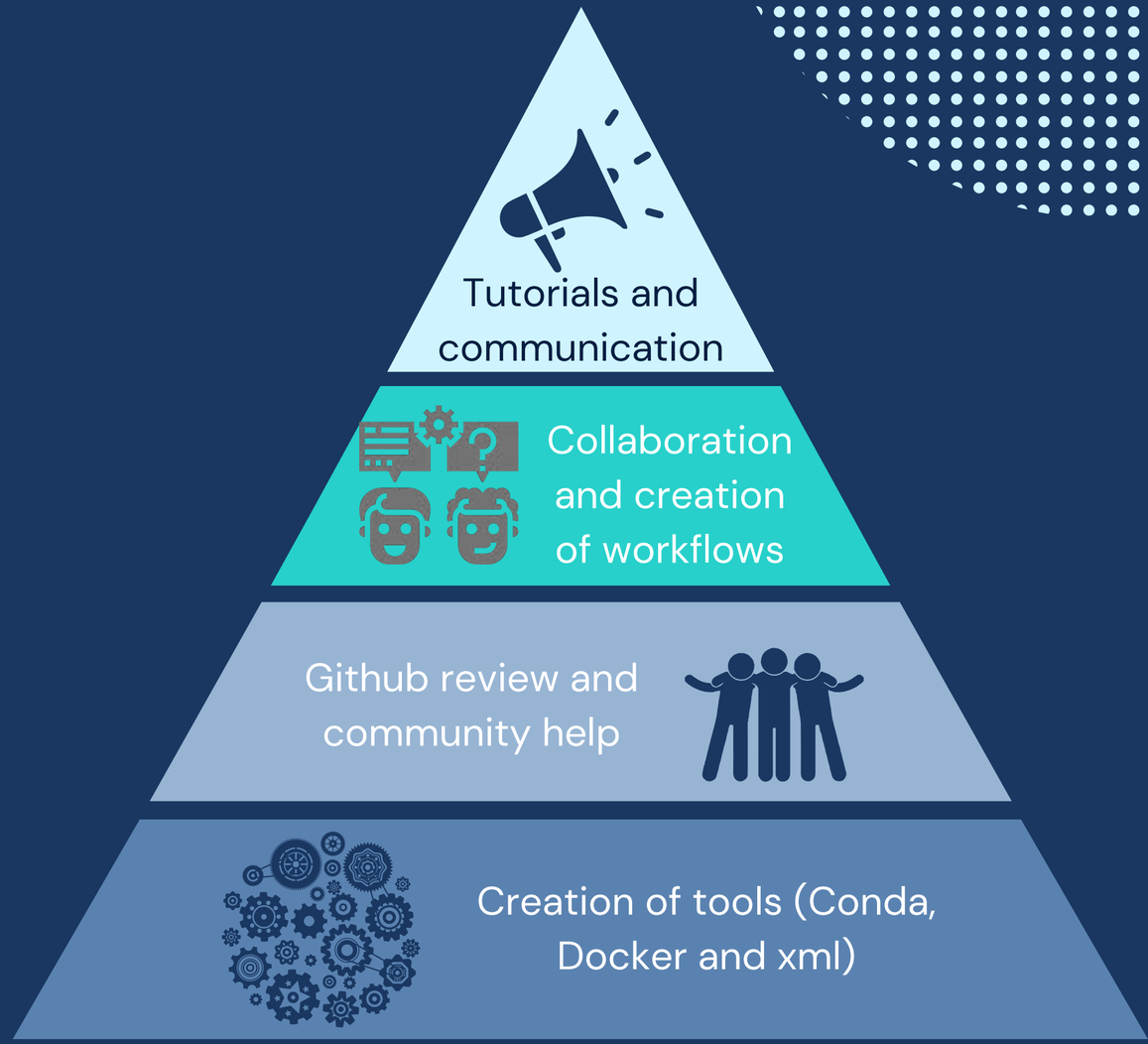
01
- Use tools from various domains (that can be plugged into workflows) through its graphical web interface.

02
- Run code in interactive environments (RStudio, Jupyter...) along with other tools or workflows.

03
- Manage data by sharing and publishing results, workflows, and visualizations.

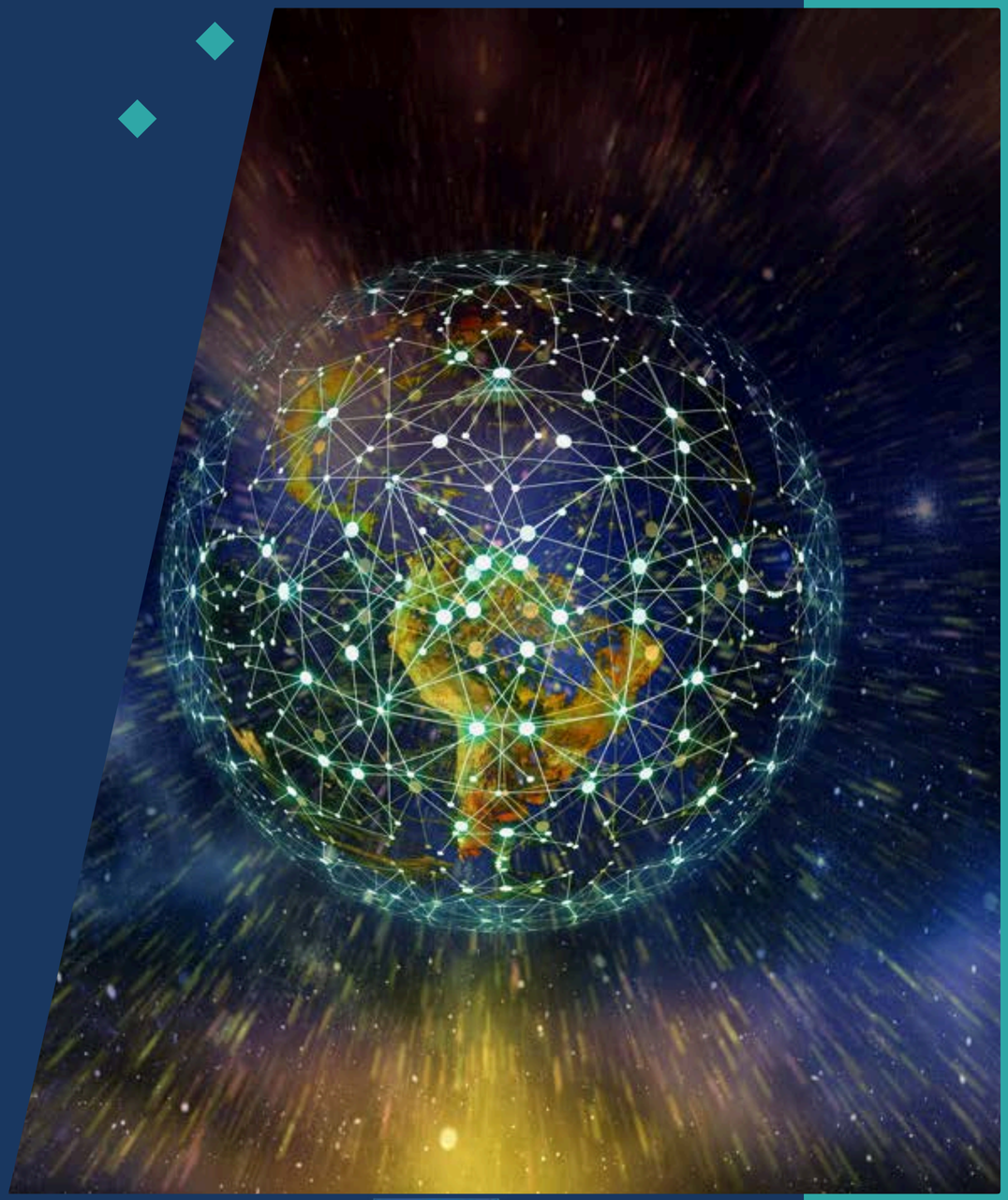
04
- Ensure reproducibility by capturing the necessary information to repeat and understand data analyses.

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



eosc | FAIR-EASE

THANK YOU



Funded by
the European Union

27-29 May 2024

International conference on Marine Data and Information Systems

SeaDataNet AISBL · Havforskningsinstituttet · Ifremer · MARIS · MEDIN · National Oceanography Centre · INGV · iMagine · IODE · eosc Blue-Cloud2026