



Open Science on Cloud using Jupyter Notebooks

# Git, Containerization, and CI/CD workflows

S. Koulouzis, G. Pelouze

(LifeWatch ERIC VLIC / MNS, University of Amsterdam)

26/06/2025



UNIVERSITY  
OF AMSTERDAM

**IR0000032 – ITINERIS, Italian Integrated Environmental Research Infrastructures System**  
(D.D. n. 130/2022 - CUP B53C22002150006) Funded by EU - Next Generation EU PNRR-  
Mission 4 "Education and Research" - Component 2: "From research to business" - Investment  
3.1: "Fund for the realisation of an integrated system of research and innovation infrastructures"



Finanziato  
dall'Unione europea  
NextGenerationEU



Ministero  
dell'Università  
e della Ricerca



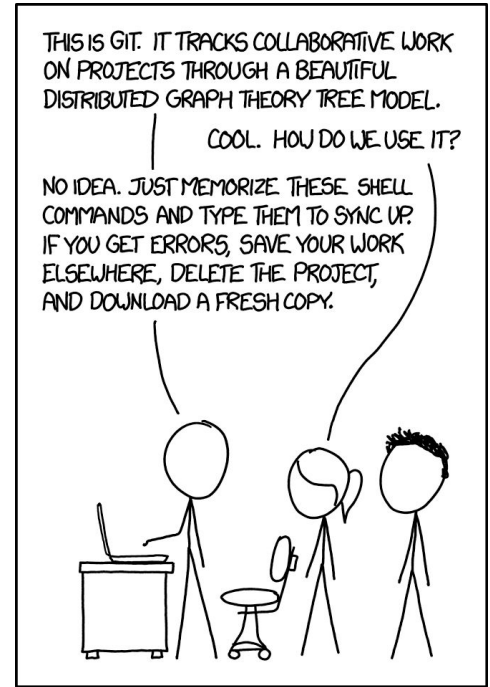
Italiadomani  
INNOVARE IL FUTURO



# Introduction to Git

# What is Git?

- An open source distributed **version control** system that helps track changes in files
- Git is the **de facto standard** for version control
- When working alone, you look at old **snapshots** of a project
- Keep a log of **who** and **when changes** were made
- Work on parallel **branches** of development
- **Github** is the most popular web platform to host Git projects
- Other web platforms are GitLab, Bitbucket



# User Story

- As an ecologist I'm writing a model for predicting the distribution of invasive species using Python
- After several iterations the folder on my laptop looks like this:

```
|— select_habitat.py
|— select_habitat_final.py
└— query.py
└— main.py
└— 2071-2100_merged_output.tif
```

# User Story

- I put my folder on GDrive to share with my colleagues
- After several iterations the the GDrive looks like this
- One year after I want to run the model with different data
- The code is no longer running because python is updated

└─ select\_habitat.py

└─ select\_habitat\_final.py

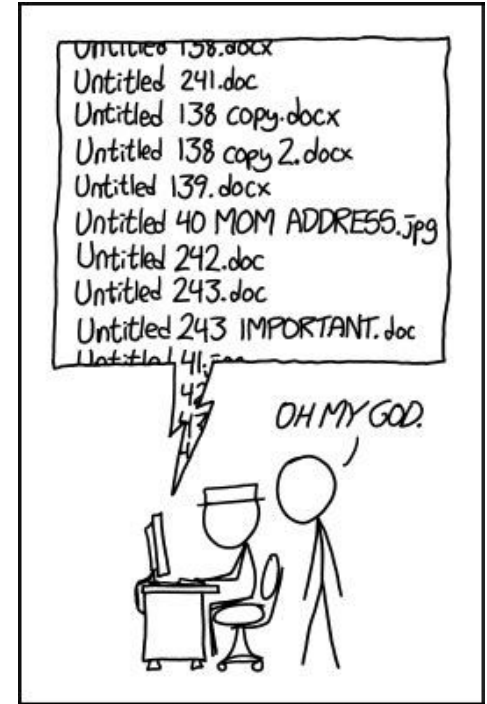
└─ query.py

└─ query\_bob.py

└─ main.py

└─ 2071-2100\_merged\_output.tif

└─ 2071-2100\_merged\_output\_bob.tif

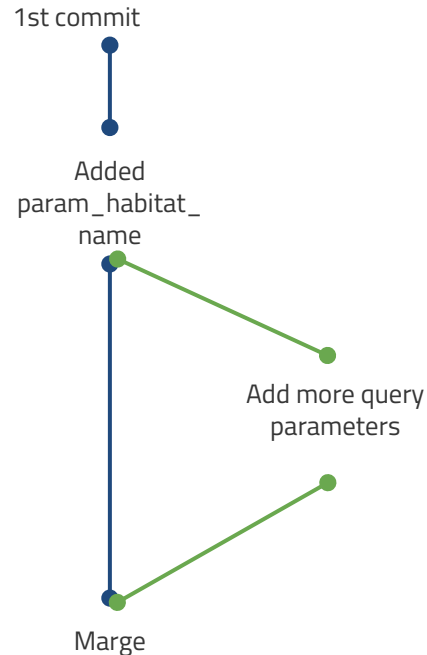


PROTIP: NEVER LOOK IN SOMEONE ELSE'S DOCUMENTS FOLDER.

# What can Git do for me ?

- Track changes in files
- Create and merge branches
- After several iterations the folder on my laptop looks like this:

```
|— select_habitat.py  
|— query.py  
|— main.py  
|— 2071-2100_merged_output.tif
```



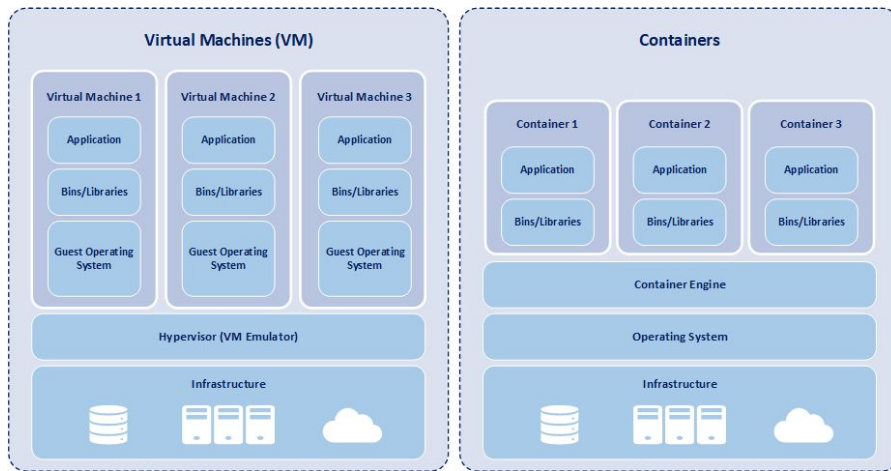
In case of fire



1. **git commit**
2. **git push**
3. **leave building**

# What is a Container and what can it do for me?

- A software that “encapsulates” code, libraries into a “container”
- The container can be saved shared and executed by any computer that supports the **containerd** standard
- Docker is the defacto technology for containers
- My code will run everywhere not just on my laptop

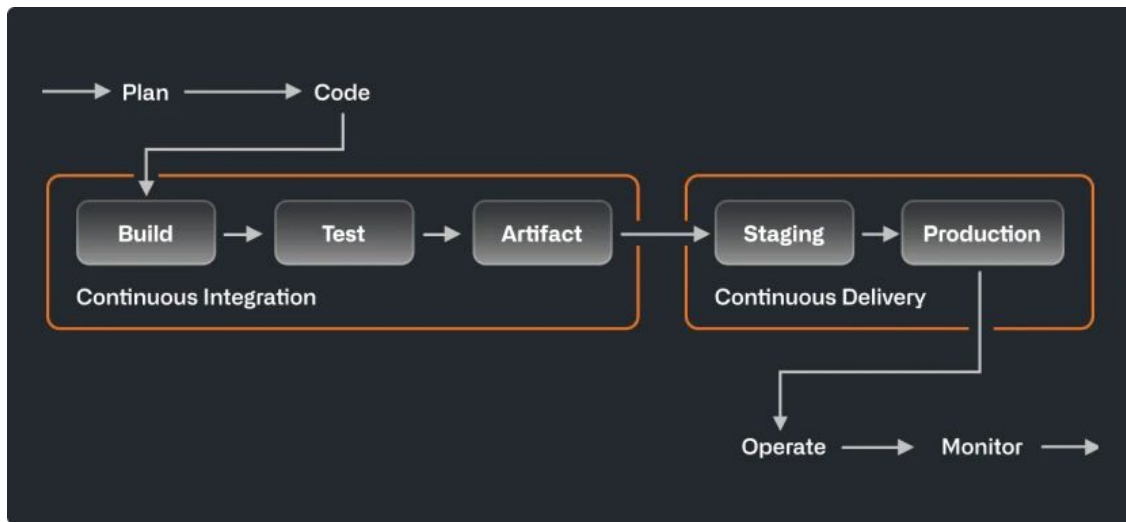


~\\_ (ツ) \\_~/~  
**IT WORKS**  
on my machine

# Introduction to CI/CD workflows

# What are CI/CD workflows?

- **Automated actions** that run when the code in a repository changes, or upon other events
  - **Continuous Integration:** build, test, various QA (spellcheck, linting, coding standards, secrets scanning, FAIRness and security assessment)
  - **Continuous Delivery:** release, deploy to testing environment
  - **Continuous Deployment:** deploy to production environments



# CI/CD workflows plateforms

- **Git + CI/CD + project managemen:** GitHub, GitLab, Bitbucket, ...
- **CI/CD only:** Argo CI, Circle CI, Jenkins, Travis, ...

The screenshot displays the GitHub Actions interface for a repository named 'NaaVRE / NaaVRE-catalogue-service'. The top navigation bar includes links for Code, Issues (1), Pull requests (1), Actions (selected), Projects, Wiki, Security (4), Insights, and Settings. A search bar is located on the right. The main content area shows a 'CI pipeline' run titled 'Merge pull request #33 from NaaVRE/add\_paas\_features #103', which is in a 'Success' state. A summary table provides key metrics: 'Triggered via push last month', 'Status: Success', 'Total duration: 1m 40s', and 'Artifacts: 1'. Below this, the 'ci-pipeline.yaml' workflow is visualized as a sequence of jobs: 'Configure workflow' (8s), 'lint / lint' (8s) and 'test / test' (46s), and 'build-docker / Build docker' (33s) followed by 'helm-helm / Build helm' (15s). A left sidebar offers a 'Summary' view and a list of jobs with expandable details.

NaaVRE / NaaVRE-catalogue-service

Code Issues 1 Pull requests 1 Actions Projects Wiki Security 4 Insights Settings

← CI pipeline

✓ Merge pull request #33 from NaaVRE/add\_paas\_features #103

Summary

Jobs

- ✓ Configure workflow
- ✓ lint
- ✓ test
- ✓ build-docker
- ✓ helm-helm

Run details

- Usage
- Workflow file

Triggered via push last month	Status	Total duration	Artifacts
gpelouze pushed 0a74c07 main	Success	1m 40s	1

ci-pipeline.yaml

on: push

```
graph LR; A[Configure workflow 8s] --> B[lint / lint 8s  
test / test 46s]; B --> C[build-docker / Build docker 33s  
helm-helm / Build helm 15s];
```

# How to create a CI/CD workflow

To create a workflow on GitHub, create a YAML file in `.github/workflows/`

The file is automatically picked-up by the actions runner and executed on the appropriate events.

```
ci-pipeline.yaml
1 name: CI pipeline
2
3 on:
4   release:
5     types: [published]
6   push:
7     branches:
8       - '**'
9
10 jobs:
11   build-docker:
12     name: Build docker
13     runs-on: ubuntu-latest
14     steps:
15       - name: Checkout
16         uses: actions/checkout@v4
17         with:
18           submodules: recursive
19
20       - name: Set up Docker Buildx
21         uses: docker/setup-buildx-action@v3
22
23       - name: Log in to the Container registry
24         uses: docker/login-action@v3
25         with:
26           registry: ghcr.io
27           username: ${ github.actor }
28           password: ${ secrets.GITHUB_TOKEN }
29
30       - name: Build docker
31         uses: docker/build-push-action@v6
32         with:
33           context: .
34           file: Dockerfile
35           tags: ghcr.io/me/my-app:latest
36           push: ${ github.event_name == 'release' }
```

**build-docker / Build docker**  
succeeded 33 minutes ago in 11s

- > ☒ Set up job
- > ☒ Checkout
- > ☒ Set up Docker Buildx
- > ☒ Log in to the Container registry
- > ☒ Build docker
- > ☒ Post Build docker
- > ☒ Post Log in to the Container registry
- > ☒ Post Set up Docker Buildx
- > ☒ Post Checkout
- > ☒ Complete job

# Hands on Git



<https://github.com/QCDIS/git-tutorial/blob/main/README.md>

# Hands on Docker



<https://github.com/QCDIS/git-tutorial/blob/main/README.md#docker>

# Hands on CI/CD workflows



<https://github.com/QCDIS/git-tutorial/blob/main/README.md#cicd-workflow>



# THANKS!



**IR0000032 – ITINERIS, Italian Integrated Environmental Research Infrastructures System**  
(D.D. n. 130/2022 - CUP B53C22002150006) Funded by EU - Next Generation EU PNRR-  
Mission 4 "Education and Research" - Component 2: "From research to business" - Investment  
3.1: "Fund for the realisation of an integrated system of research and innovation infrastructures"

