ITINERIS

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

LifeWatch ERIC

University of Amsterdam

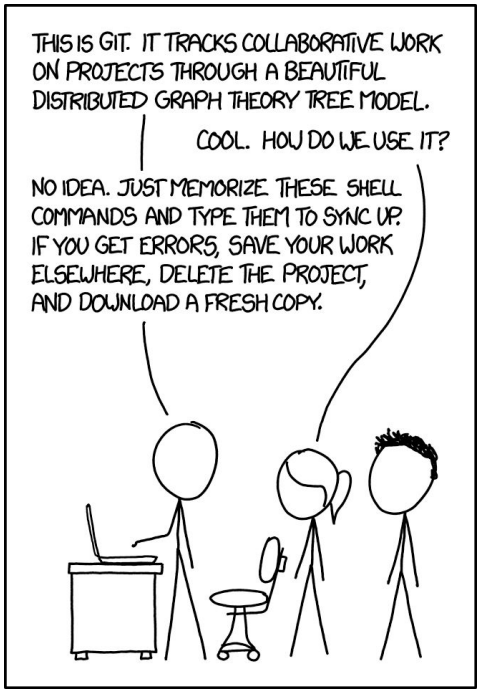
IR0000032 – ITINERIS, Italian Integrated Environmental Research Infrastructures System
(D.D. n. 130/2022 - CUP B53C22002150006) Funded by EU - Next Generation EU PNRR-Mission 4 "Education and Research" - Component 2: "From research to business" - Investment 3.1: "Fund for the realisation of an integrated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

Finanziato dall'Unione europea NextGenerationEU
Ministero dell'Università e della Ricerca
Italiadomani

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 spaces 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

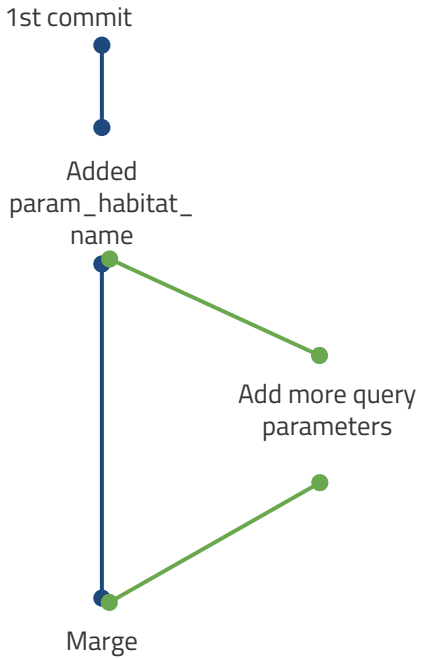
- I put my folder on GDrive to share with my colleagues
- After several iterations the the GDrive looks like this
- One year after I want to run the model with different data
- The code is no longer running because python is updated

```
├── 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
```

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
```

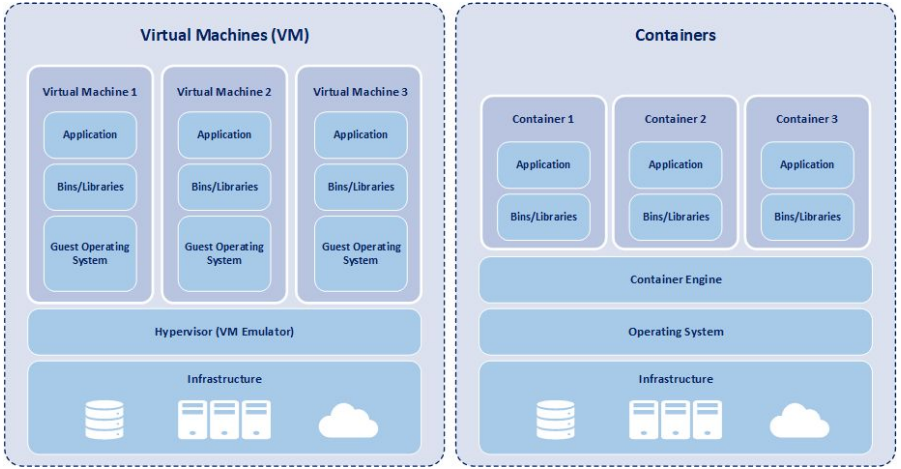


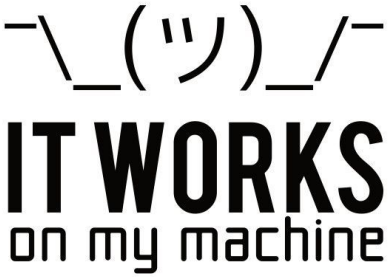
In case of fire

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 **continerd** standard
- Docker is the defacto technology for containers
- My code will run everywhere not just on my laptop

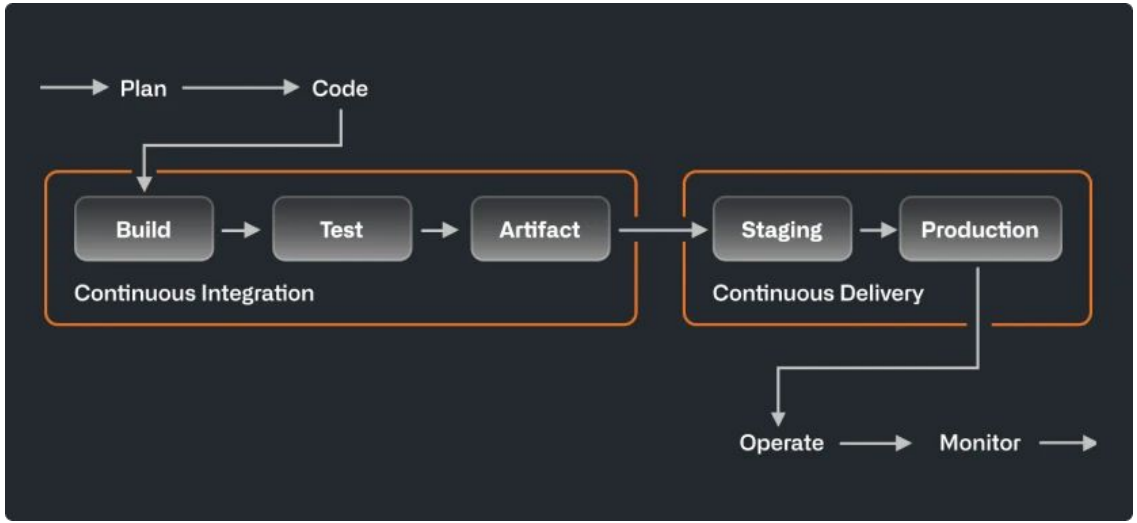




Introduction to CI/CD workflows

What are CI/CD workflows?

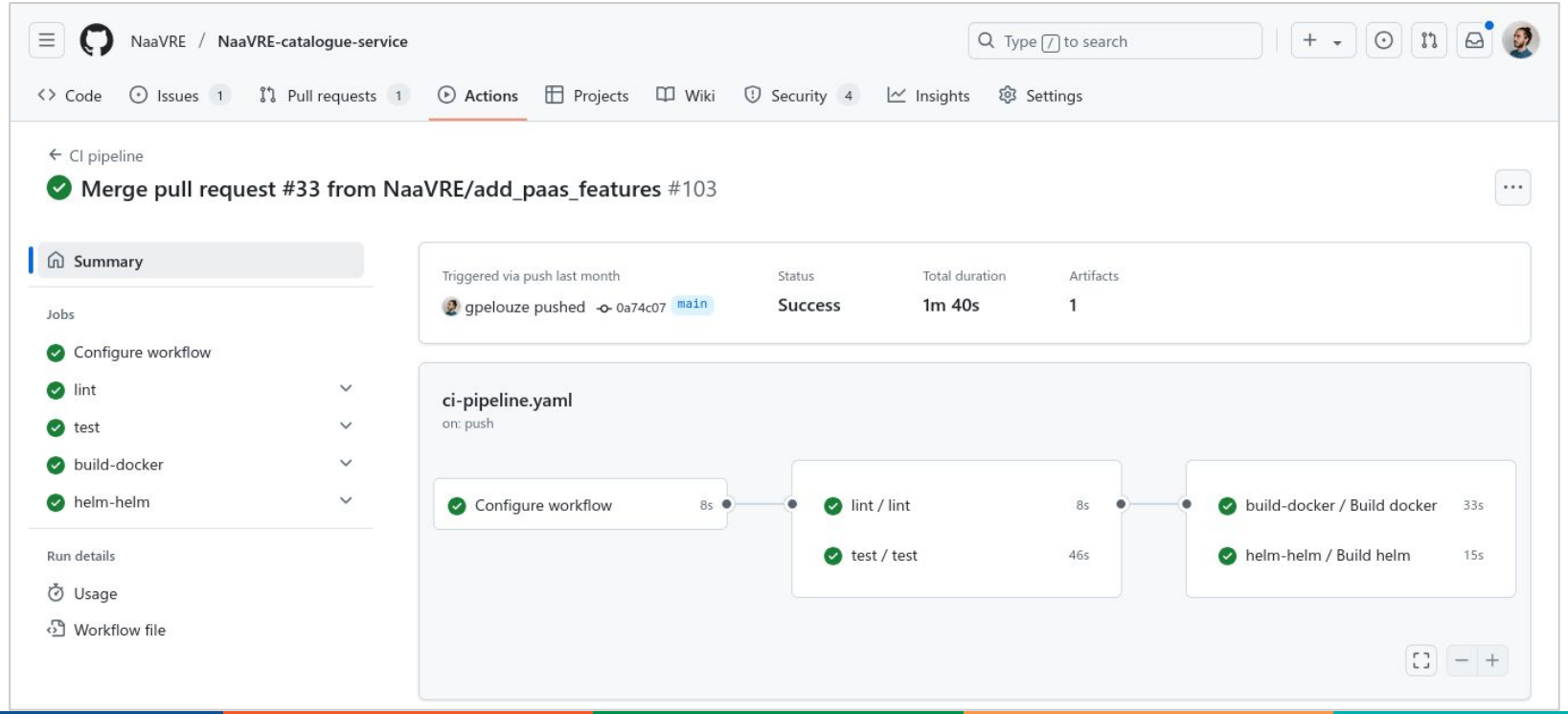
- **Automated actions** that run when the code in a repository changes, or upon other events
  - **Continuous Integration:** build, test, various QA (spellcheck, linting, coding standards, secrets scanning, FAIRness and security assessment)
  - **Continuous Delivery:** release, deploy to testing environment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



https://github.com/resources/articles/devops/ci-cd

CI/CD workflows plateforms

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



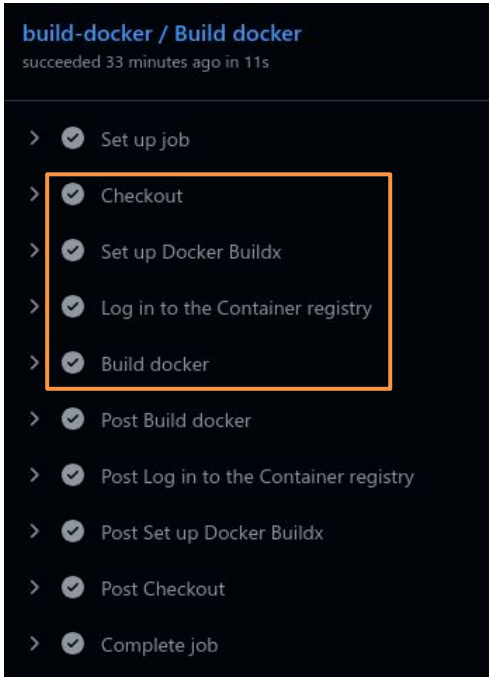
How to create a CI/CD workflow

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 — Raw

```yaml
name: CI pipeline

on:
  release:
    types: [published]
  push:
    branches:
      - '**'

jobs:
  build-docker:
    name: Build docker
    runs-on: ubuntu-latest
    steps:
      - name: Checkout
        uses: actions/checkout@v4
        with:
          submodules: recursive

      - name: Set up Docker Buildx
        uses: docker/setup-buildx-action@v3

      - name: Log in to the Container registry
        uses: docker/login-action@v3
        with:
          registry: ghcr.io
          username: ${{ github.actor }}
          password: ${{ secrets.GITHUB_TOKEN }}

      - name: Build docker
        uses: docker/build-push-action@v6
        with:
          context: .
          file: Dockerfile
          tags: ghcr.io/me/my-app:latest
          push: ${{ github.event_name == 'release' }}
```



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

ITINERIS

THANKS!

LifeWatch ERIC

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