FAIR_bioinfo for bioinformaticians

Introduction to the tools of reproducibility in bioinformatics

C. Hernandez[1]   T. Denecker[1]   J.Sellier[2]   C. Toffano-Nioche[1]

[1]Institute for Integrative Biology of the Cell (I2BC)
UMR 9198, Université Paris-Sud, CNRS, CEA
91190 - Gif-sur-Yvette, France

[2]Institut Français de Bioinformatique
à compléter

Sept. 2020

Conclusion

Training schedule

Day 1:

- Introduction to FAIR_bioinfo
- Encapsulation (docker)
- Workflow (SNAKEMAKE)
- IFB resources (, slurm)

Day 2:

- History management (git, GitHub)
- Software environment management (CONDA)
- Traceability with notebooks (jupyter,)
- Sharing and disseminating (GitHub, zenodo)

Let's take a step back.

FAIR_bioinfo

Findable

Easy to find
protocols
(GitHub, GitHub Pages)
with DOI (zenodo)

Accessible

Open source
(GitHub,
docker,
CONDA, …)

Interoperable

Think "workflow"
(SNAKEMAKE +
docker / CONDA)
locally or on
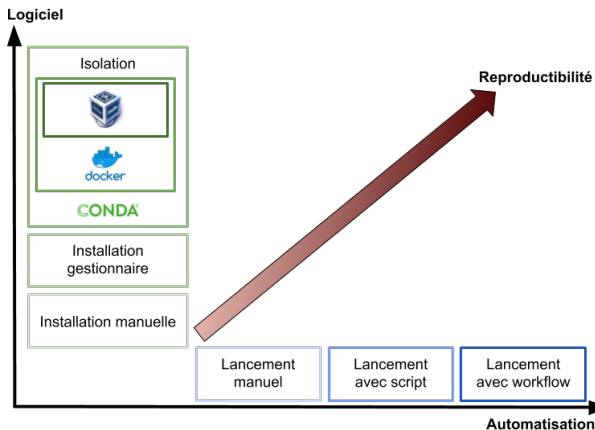servers (S, slurm)

Reusable

Replayable
protocols
(jupyter, 🔴) in virtual
environments
(docker / CONDA)

FAIR_bioinfo

Reproducibility is a multi-dimensional process

So... What now?

What now?

Automation

- Manual
- Write a script
- Use a workflow manager

Software

- Local installation
- Package manager
- Conda environment
- Image / container
- Virtual machine

Continuous integration