

FAIR_bioinfo : Open Science and FAIR principles in a bioinformatics project

How to make a bioinformatics project more reproducible

C. Hernandez¹ T. Denecker² J. Seiler² G. Le Corguillé²
C. Toffano-Nioche¹

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

²IFB Core Cluster taskforce

June 2021



General information

Practical information:

- Dates: June 28th - 30th
- Location: Institut des Systèmes Complexes, 113 rue Nationale, 75013-Paris
- Courses: 9:00 to 17:30
- Meal: 12:30-14:00
- Pauses: 10:30-11:00 + 15:30-16:00
- 2 days of courses + 1 day of course building

Round table:




- Teachers
- Learners

Ressources:








- 
- GitLab
- \LaTeX

Training schedule

Day 1:

- Introduction to reproducibility
- History management (3 Practical Sessions,  **git**,  **GitHub**)
- Control your development environment (1 PS, **CONDA**)
- Encapsulation (2 PS,  **docker**)

Day 2:

- Workflow (2 PS,  **SNAKEMAKE**)
- Traceability with notebooks (2 PS,  **jupyter**, )
- IFB resources (2 PS,  **slurm**, )
- Sharing and disseminating ( **GitHub**,  **zenodo**)
- Conclusion

Day 3:

- Empowerment and improvement of resources

Table of contents

- 1 Introduction to reproducibility
- 2 History management
- 3 Control your development environment
- 4 Workflow
- 5 Tracability with Notebook
- 6 IFB resources
- 7 Sharing and dissemination
- 8 Conclusion
- 9 3rd Day

Use cases

Reproducibility - how far and for what?



Do it all!



Workflow



Environment



Code





























nbis-reproducible-research.readthedocs.io/en/latest

Reproducibility tools

Coloured dots: symbolise tools

Variable intensities: according to expertise
not needed, basic use, advanced use (expert)

	  	clone branch
	  	collaborate pages (actions)
	  	install yml (recipe)
 , 	  	run dockerfile
	  	run snakemakefile
 , 	  	run notebookfile
	  	download submit

Use cases

What we learned during the courses



Day 1

Day 2

Day 1: Git (advanced), Github (basic), conda (advanced), zenodo (basic), docker (advanced)
Day 2: snakemake (advanced), jupyter (advanced), github (advanced)

Brainstorming

Choose coloured dots for tools to represent:

- Daily use
- Share an analysis with colleagues or your supervisor (they may want to perform)
- A part of the material and method for a publication