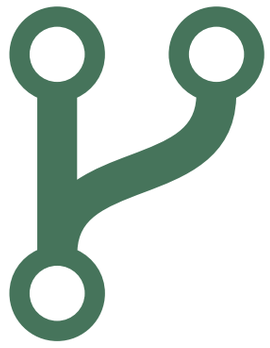
Best practices in scientific programming

What can we learn from software engineering?

Reproducibility



Version control

Documentation

Style

Open source

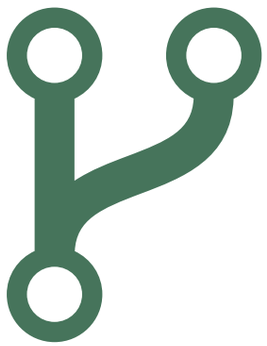
Unit testing

Virtual environments

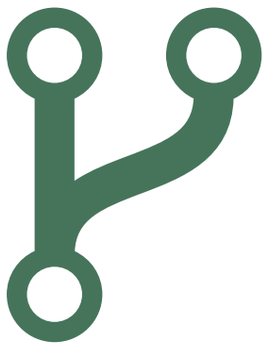
Code review

Reproducibility

- Write reproducible scripts and pipelines
- Avoid "semi-scripting"
- Not just helpful for others, but also for your future self

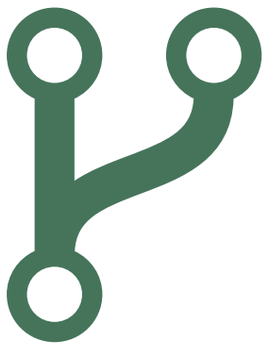


Version control



Version control

```
analysis_script.py
```



Version control

analysis_script_final_v2_New+DavidesComments_versionC.py

Version control

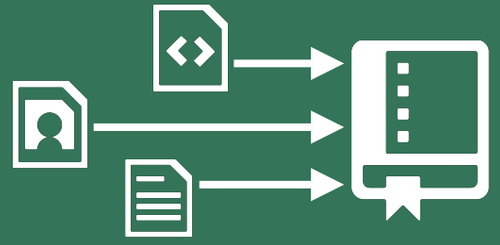
Repository

Commit

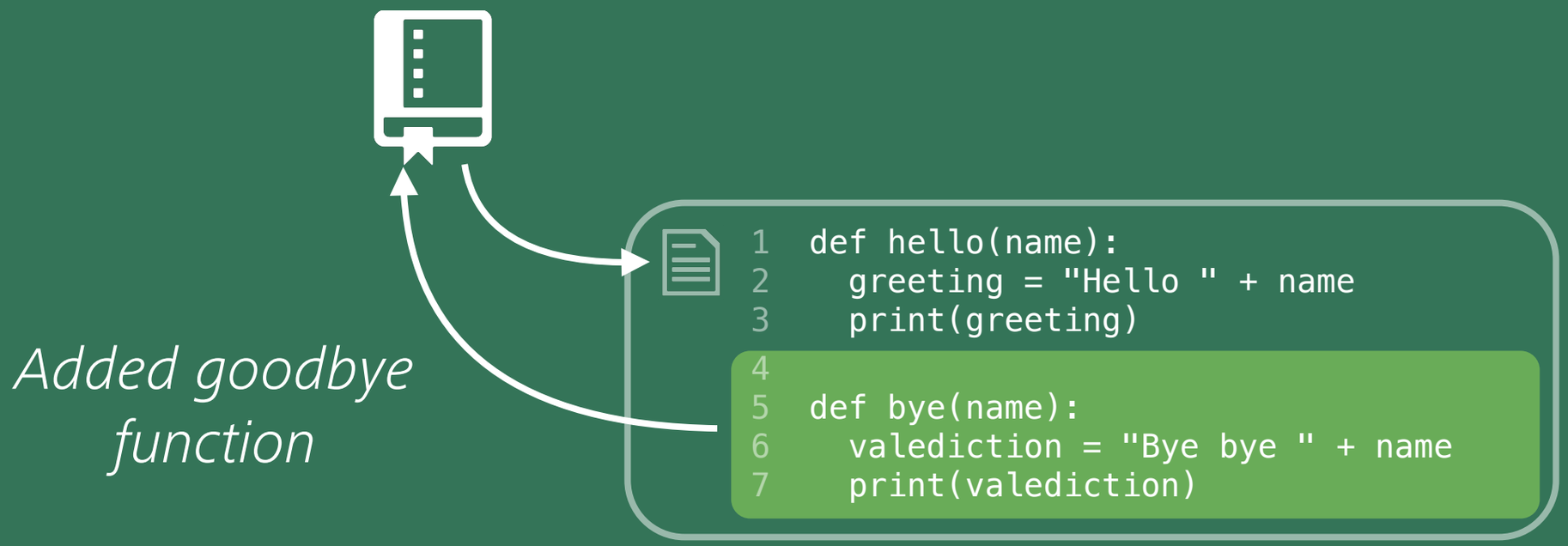
Branch

Fork

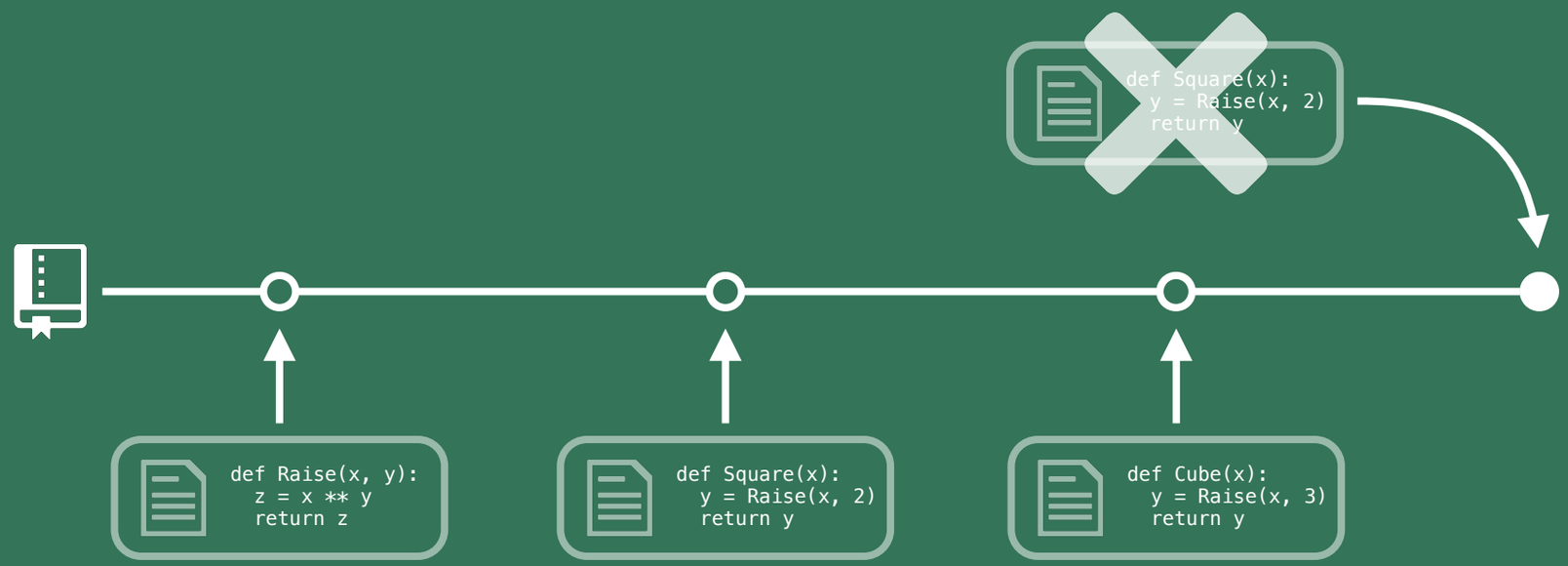
Repository



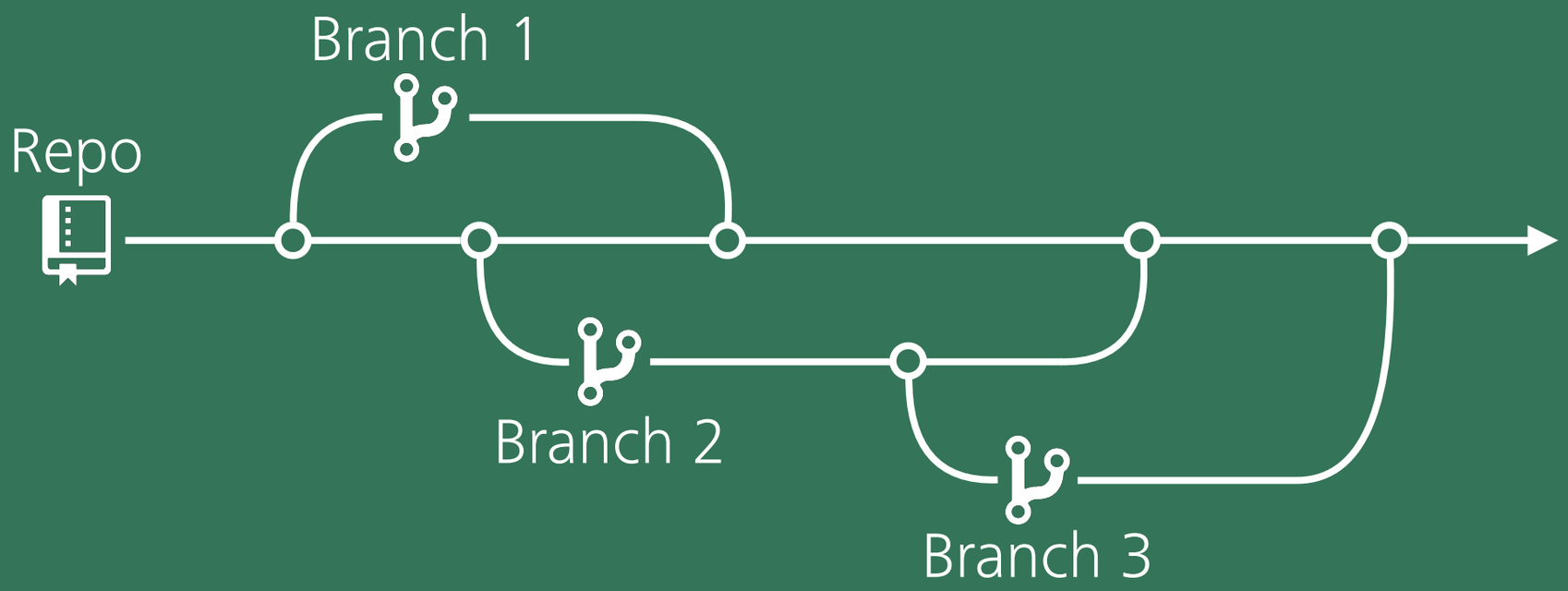
Commit


1 def hello(name):
2 greeting = "Hello " + name
3 print(greeting)
4
5 def bye(name):
6 valediction = "Bye bye " + name
7 print(valediction)
Added goodbye function

The timeline



Branching



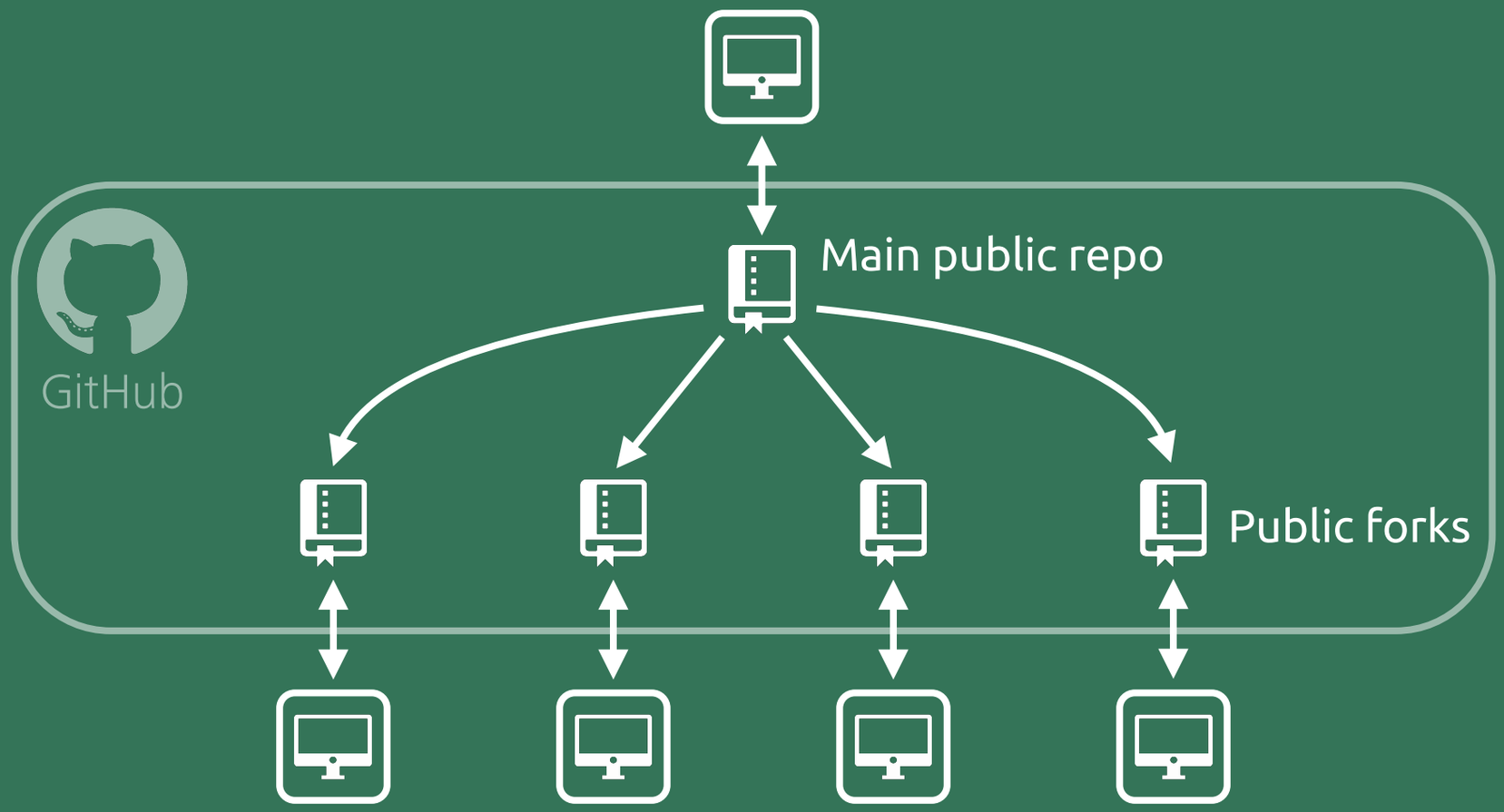
Branching

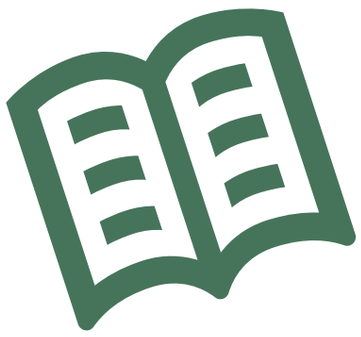
Develop new features or explore new ideas

Keep a development version separate from a stable version

Easier to deal with multiple collaborators on a single repo

Forking





Documentation

- Use clear variable/function names
- Comment your code as appropriate
- Write a readme file to help people get started with your project

Style

- Try to follow the conventions of the language you are writing in
- Use a code formatter (e.g. Black in Python)

Open source

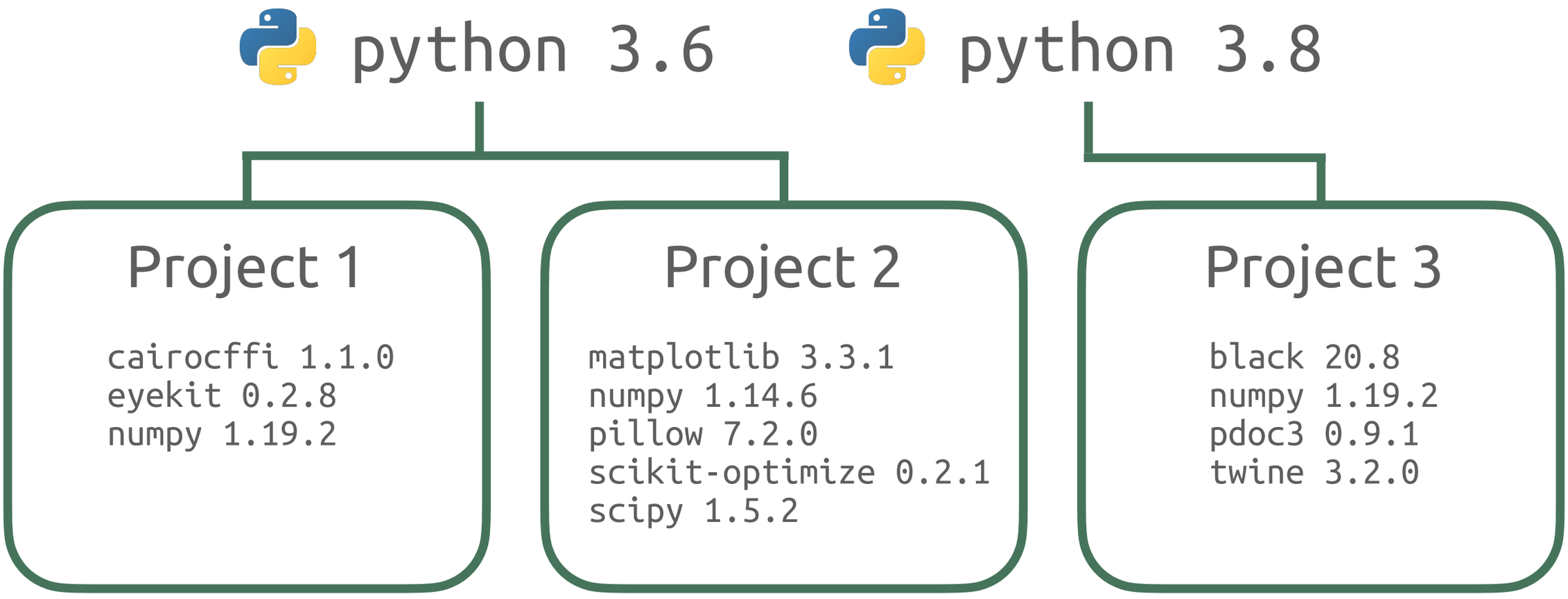
- Prefer open-source software if possible
- Open code & data: Work transparently and openly right from the start
- Give back to the community if you can

Unit testing

- Write modular code
- Write tests for your code
- Run tests frequently to check things are still working

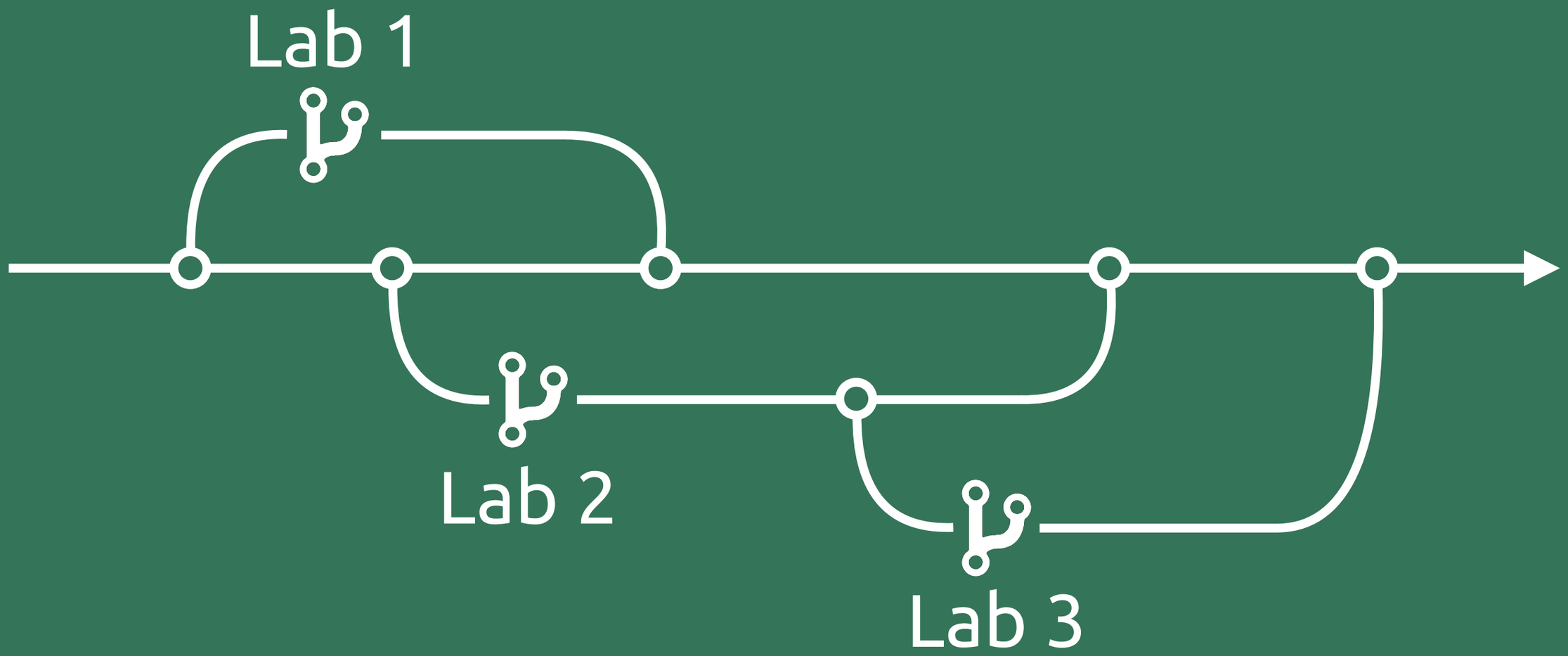
Virtual environments

- Find out how to manage multiple copies of Python and how to create virtual environments within your setup
- Use separate environments for each project, and document the package versions (e.g. pip freeze)


python 3.6
python 3.8
Project 1
cairocffi 1.1.0
eyekit 0.2.8
numpy 1.19.2
Project 2
matplotlib 3.3.1
numpy 1.14.6
pillow 7.2.0
scikit-optimize 0.2.1
scipy 1.5.2
Project 3
black 20.8
numpy 1.19.2
pdoc3 0.9.1
twine 3.2.0

Code review

- Regularly review your code and refactor if necessary
- Meet up with friends and review each other's code


Lab 1
Lab 2
Lab 3

Grazie mille

Questions, comments, tips...?