Introduction to Unix for bioinformatics

Applied bioinformatics (BBT045)

Agata Marchi

January 23rd, 2025

What is Unix?

- Operating system developed in the 1970s.
- Backbone of many modern UNIX-like operating systems (Linux, macOS).
- Provides simple but powerful tools that can be combined to perform complex tasks.
- Has only a command line interface.

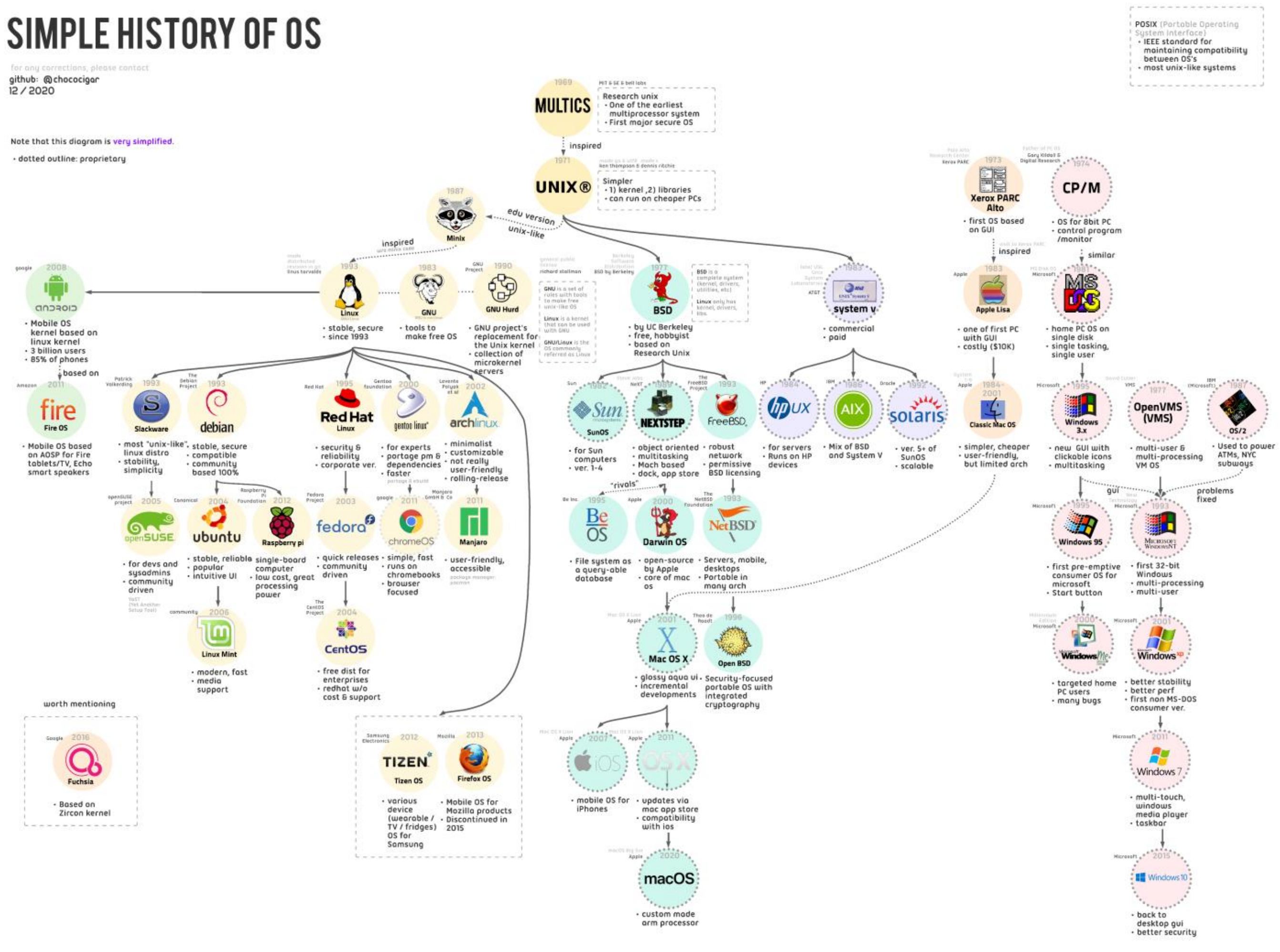


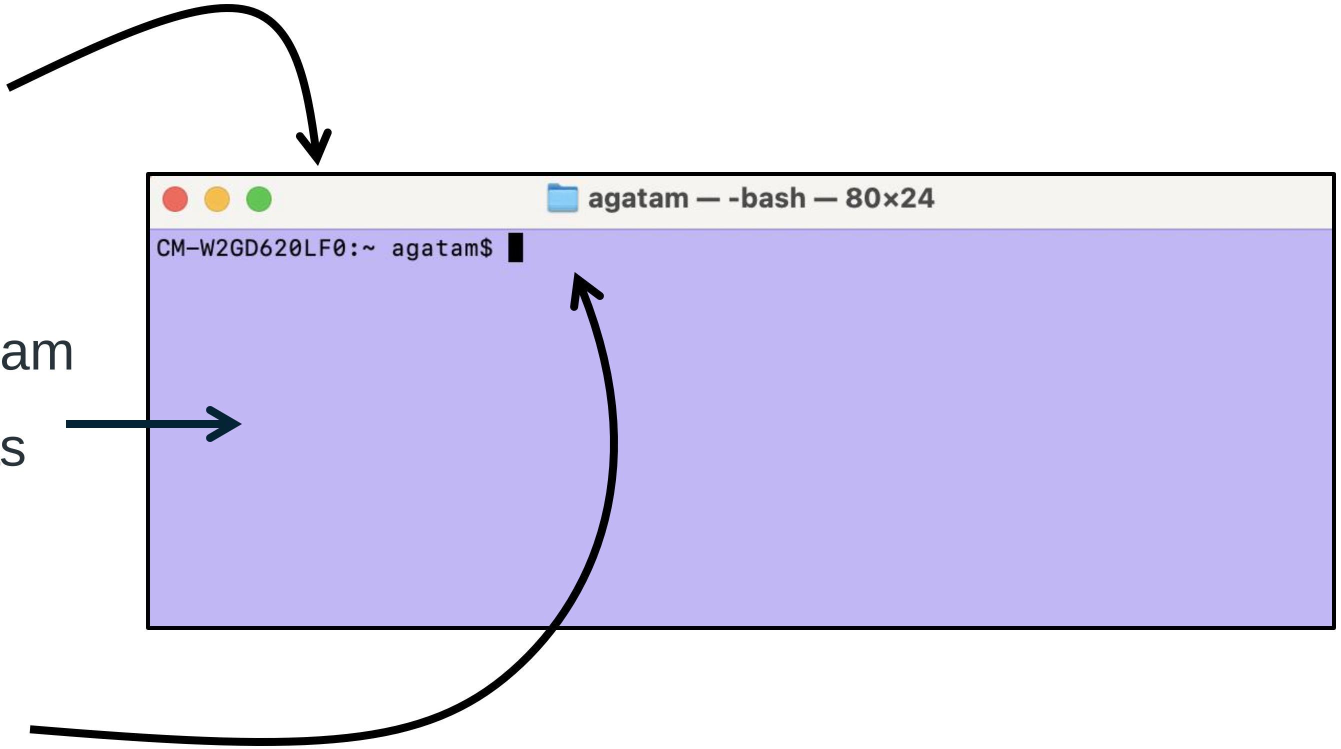
Image taken from here

The Unix philosophy

1. Programs should do one thing and do it well.
2. Programs should work together.
3. Programs should handle text because that is a universal interface.

Some terminology

- **Terminal:** program that displays a graphical interface and allows you to interact with the shell.
- **Shell:** command line interpreter program that processes commands and outputs the results.
- **Command line:** where you type your commands.

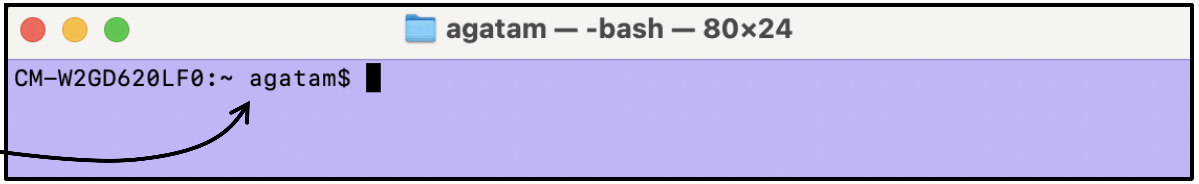


```
CM-W2GD620LF0:~ agatam$ 
```

Terminal, shell and command line are not the same thing, but these terms are often used interchangeably.

Some more terminology

Prompt: text after which you enter the commands (usually ends with $)



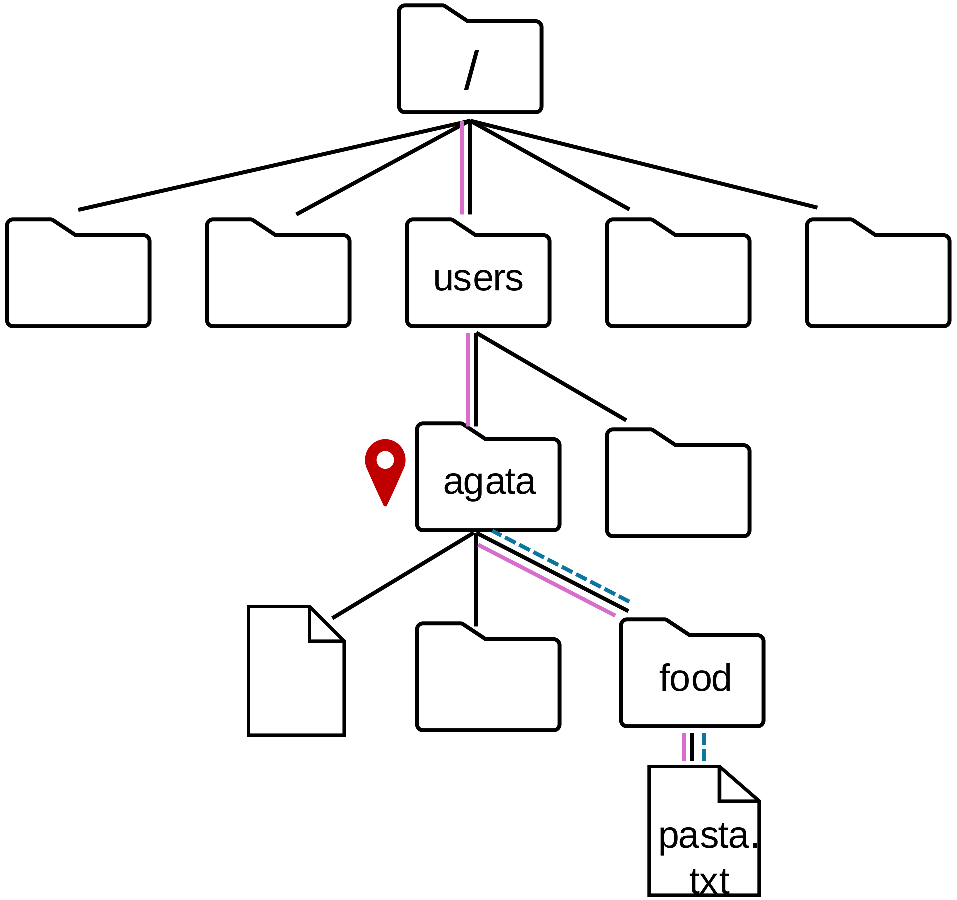
```
CM-W2GD620LF0:~ agatam$ 
```

Absolute path: the complete route to a file or directory from the system’s root.

/users/agata/food/pasta.txt

Relative path: the location of a file or directory in relation to the current working directory.

food/pasta.txt



Why Unix for bioinformatics?

- Complex data analysis by chaining simple commands

  → No need to write intermediate files.

- Automating repetitive tasks.
- Tools are optimized to work with large text files.
- High-performance computer clusters runs Unix.
- A lot of bioinformatics tools are tailored to run in a Unix environment and only come with a command line interface.

Tutorial and homework

https://bengtssonpalme.github.io/MPBIO-BBT045-2025/

Tips:

- Keep your files organized → Make directories for different things you are working on.
- Name files so you understand what they contain.
- Don’t use spaces in files and directories names.