Linux command line and nf-core introduction

Edoardo Piombo

Let' download stuff

Windows people should do the following:
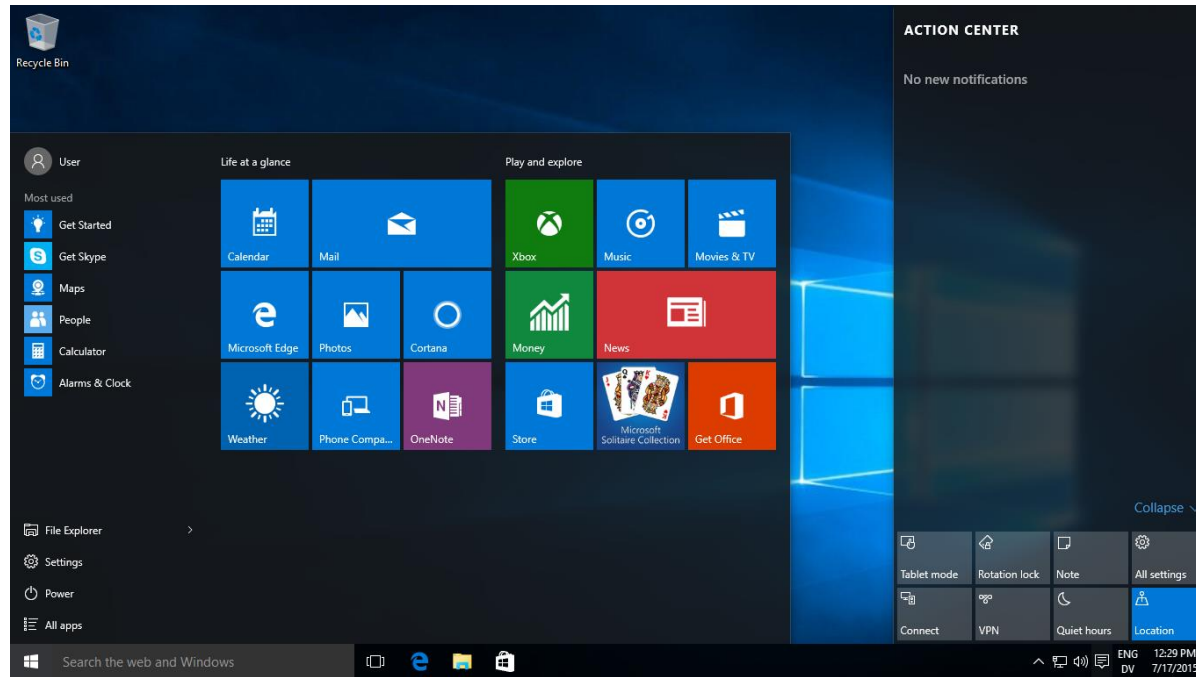
- download moba X term from here: https://mobaxterm.mobatek.net/

- Do "new session"

- Choose "shell"

- Choose "bash"

- Enter "open /home/mobaxterm" to open the directory in the file explorer

People using Macs or Linux just need to open a terminal:

- Mac: https://support.apple.com/guide/terminal/open-or-quit-terminal-apd5265185d-f365-44cb-8b09-71a064a42125/mac

- Linux: https://ubuntu.com/tutorials/command-line-for-beginners#3-opening-a-terminal

Umeå Plant
Science Centre

There are different ways of interacting with computers

Graphical user interface (GUI)

Command-line (or text) interface (CLI). Also called shell

Linux

- Operating system (OS)

- Microsoft Windows, MacOS, and **Linux**

- Linux – Versatile, open source, free,
    used on supercomputers.

- The most common way to run complex bioinformatic
    analyses, on supercomputers or your own pc,
    is to use the **Bash shell.**

Umeå Plant Science Centre
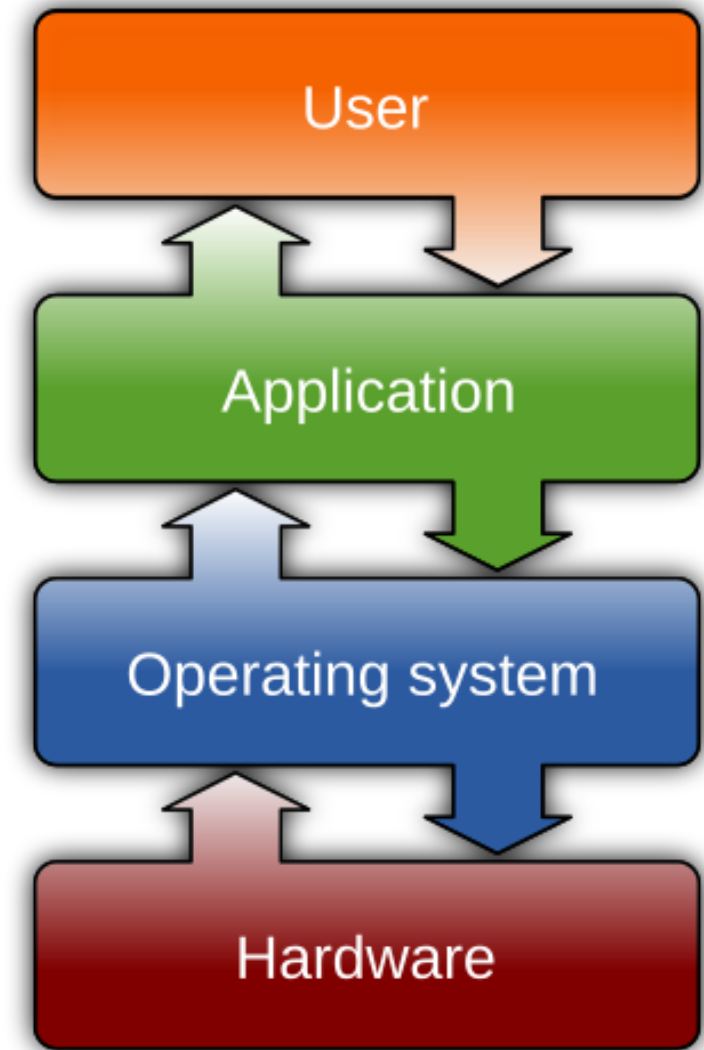
What we will do today

- Tutorial about the bash shell

Break

- Learning about **nf-core**

- Exercise about **nf-core**

Umeå Plant
Science Centre

Open a terminal: Mac

- Open launchpad

- Write "terminal" in the search bar.

- Click on Terminal

Umeå Plant
Science Centre
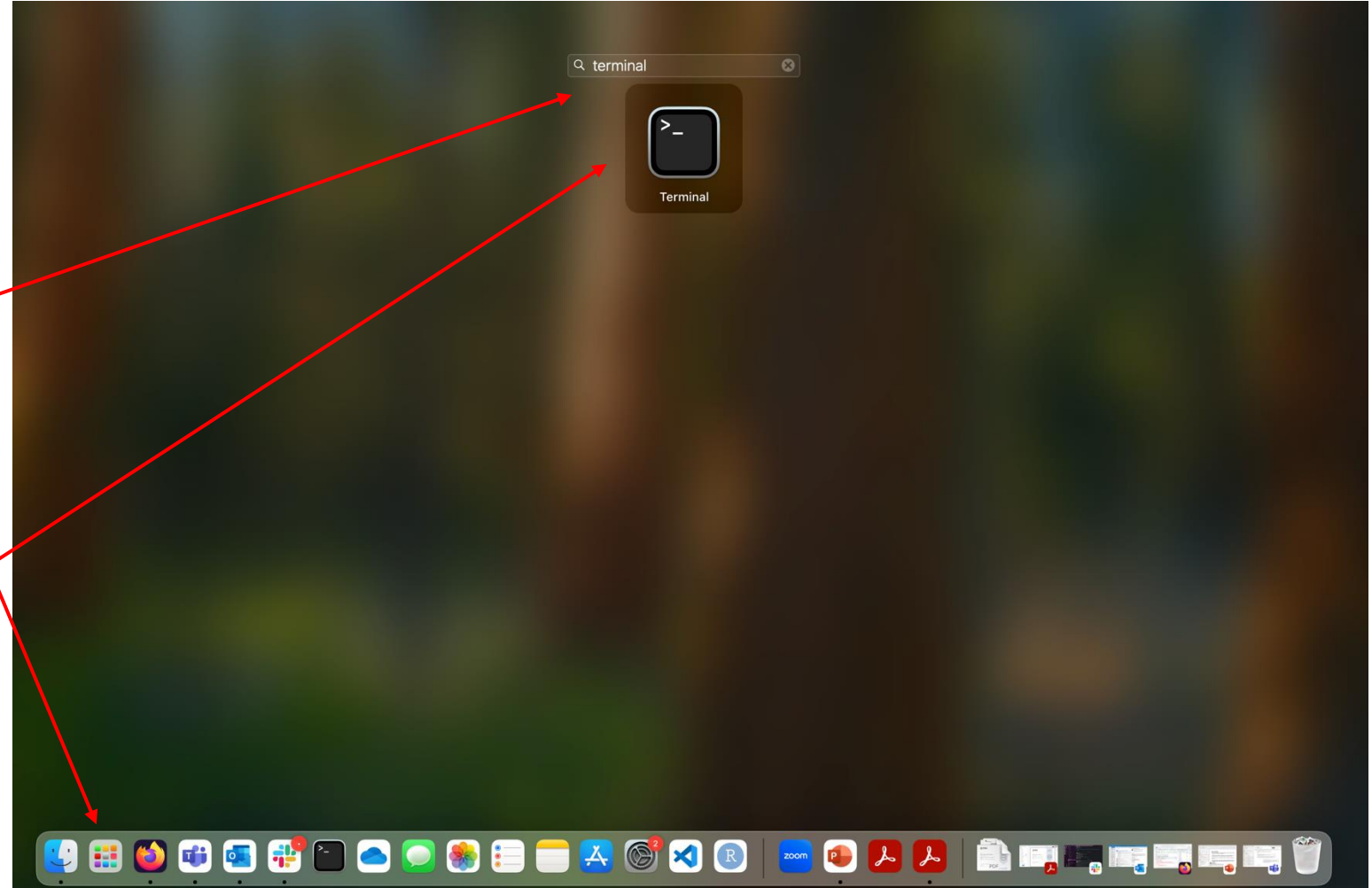
Open a terminal: Mac

- You should see something like this

```
edoardo — -zsh — 80×24
Last login: Wed May 21 08:53:53 on console
(base) edoardo@client227-209 ~ %
```

Umeå Plant
Science Centre

Open a terminal: Windows

- Open MobaXterm

- Start a new session

Open a terminal: Windows

- Open MobaXterm

- Start a new session

- Select "shell"

Umeå Plant
Science Centre

Open a terminal: Windows

- Open MobaXterm

- Start a new session

- Select "shell"

- Select an empty startup directory or make a new one

Umeå Plant Science Centre

Open a terminal: Windows

- Open MobaXterm

- Start a new session

- Select "shell"

- Select an empty startup directory or make a new one

Umeå Plant
Science Centre

Open a terminal: Windows

- Open MobaXterm

- Start a new session

- Select "shell"

- Select an empty startup directory or make a new one

Umeå Plant Science Centre

Open a terminal: Windows

- Open MobaXterm

- Start a new session

- Select "shell"

- Select an empty startup directory or make a new one

- Click "OK" to open a terminal

Session settings

| SSH | Telnet | Rsh | Xdmcp | RDP | VNC | FTP | SFTP | Serial | File | Shell | Browser | Mosh | Aws S3 | WSL |

Basic Shell settings

Terminal shell [Bash ▾]   Startup directory C:\Users\Elena van Zalen\One 📁

Advanced Shell settings   Terminal settings   ★ Bookmark settings

Local shell session

OK   Cancel

Umeå Plant Science Centre
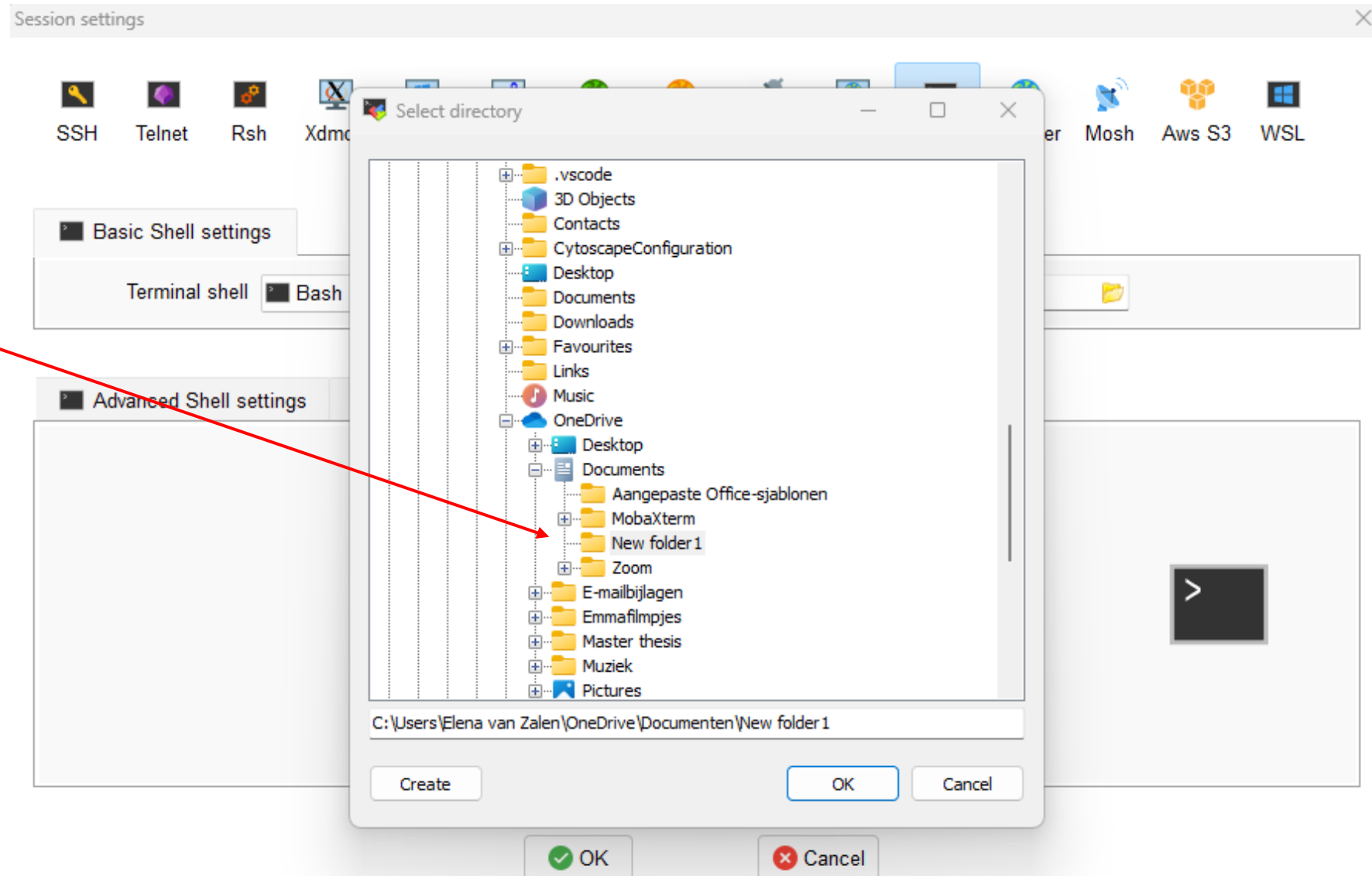
Open a terminal: Windows

- Open MobaXterm

- Start a new session

- Select "shell"
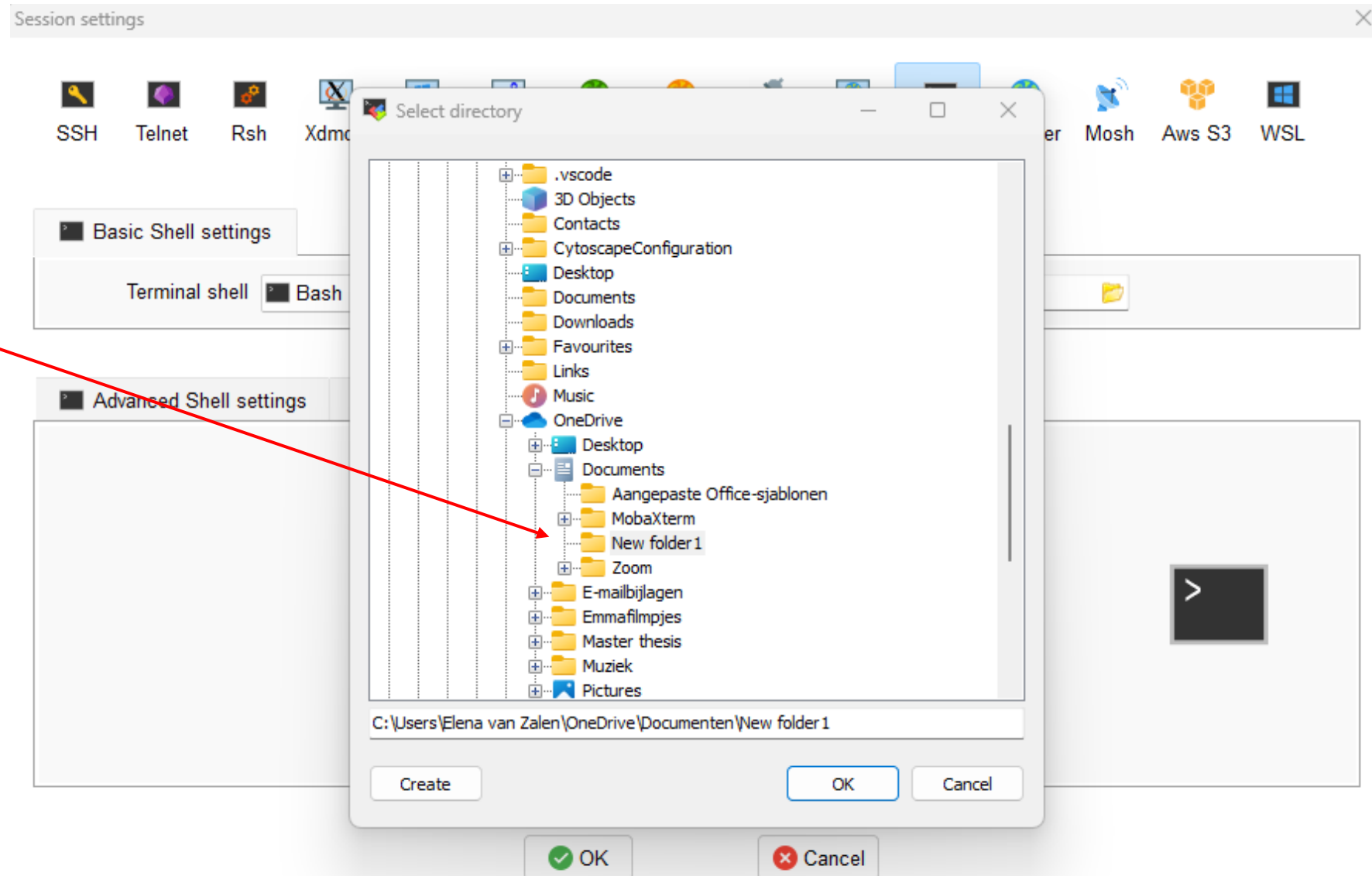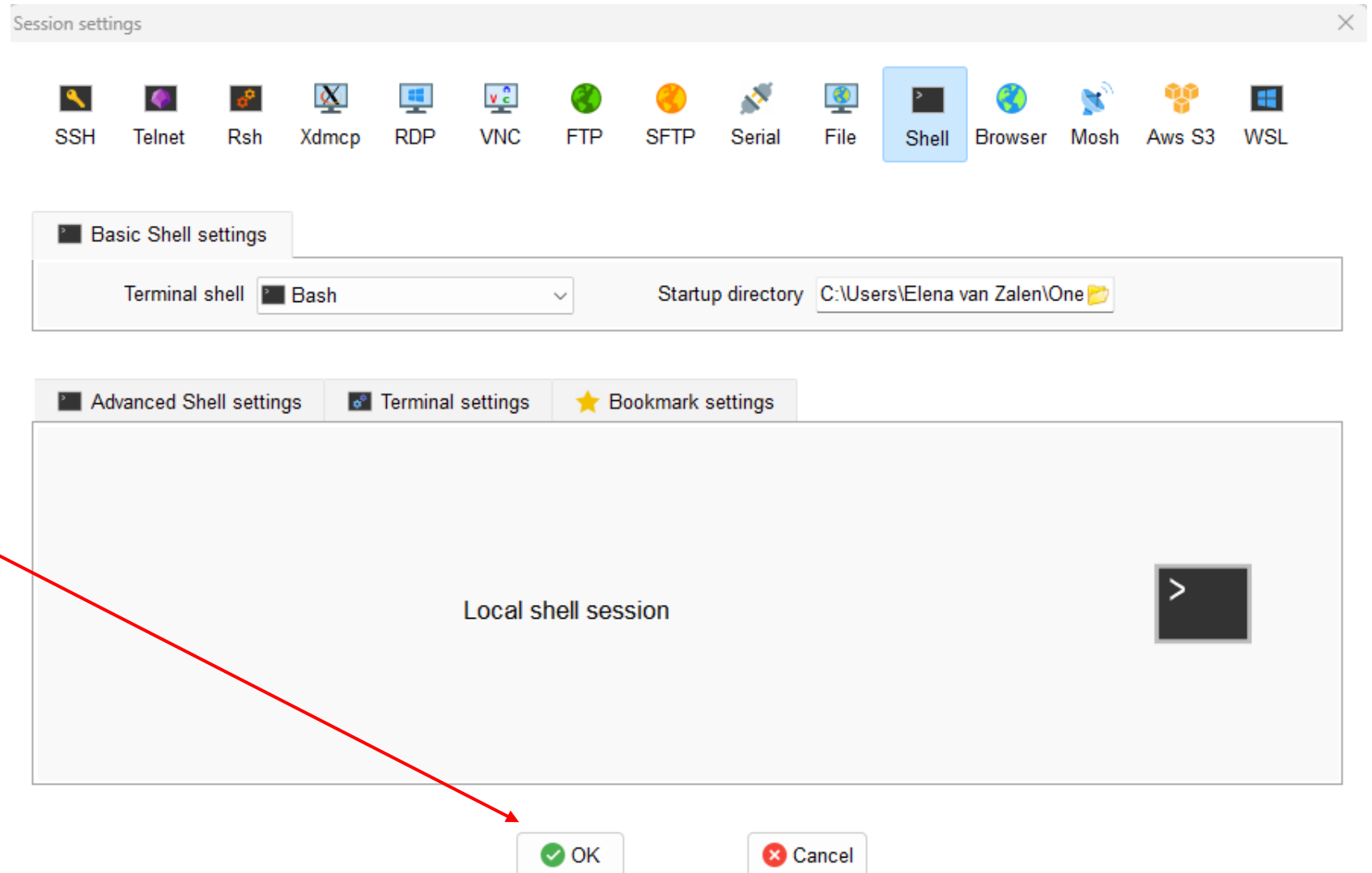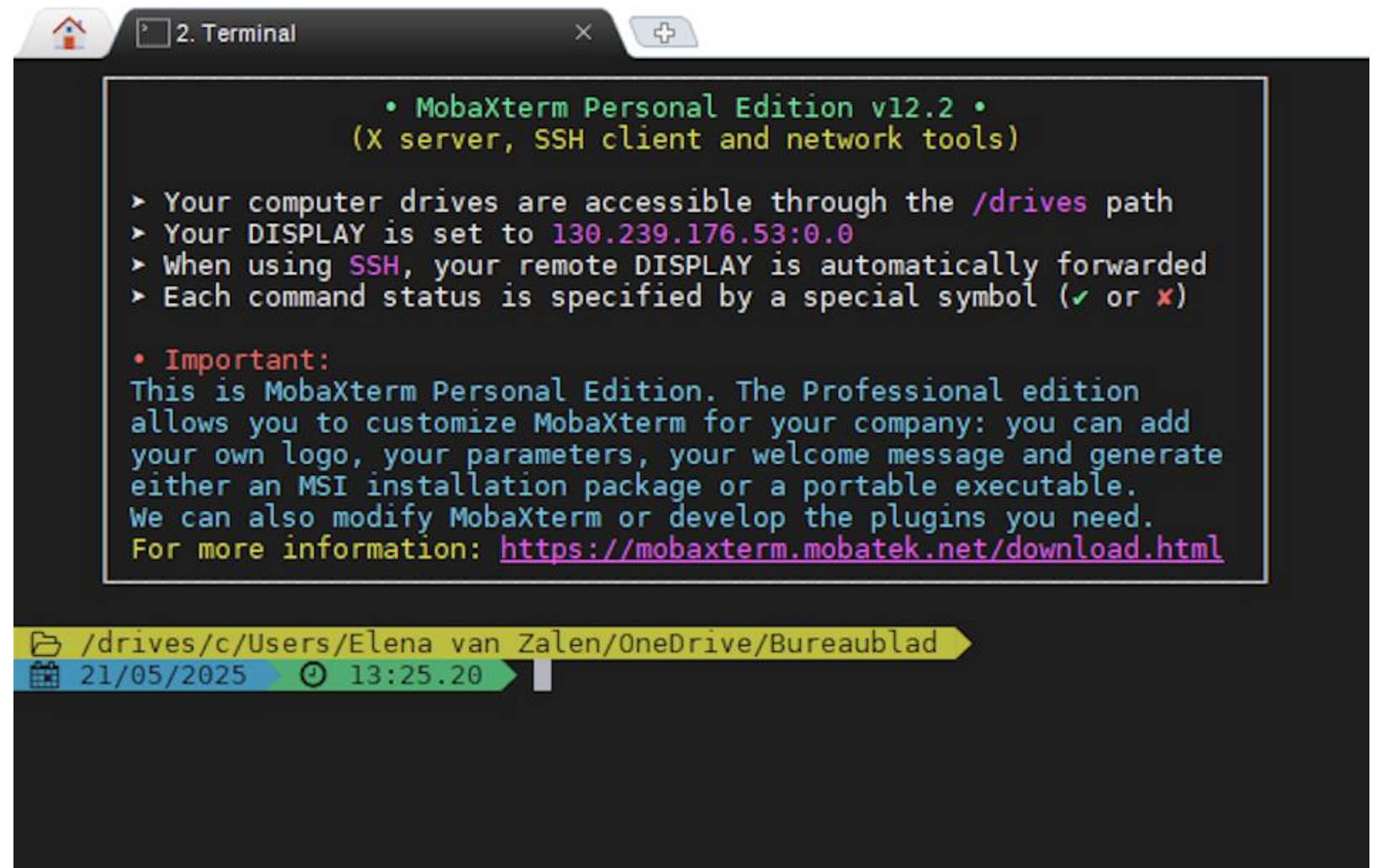
- Select an empty startup directory or make a new one

- Click "OK" to open a terminal

- You should see something like this


```
                    • MobaXterm Personal Edition v12.2 •
                   (X server, SSH client and network tools)

  ➤ Your computer drives are accessible through the /drives path
  ➤ Your DISPLAY is set to 130.239.176.53:0.0
  ➤ When using SSH, your remote DISPLAY is automatically forwarded
  ➤ Each command status is specified by a special symbol (✔ or ✘)


  • Important:
  This is MobaXterm Personal Edition. The Professional edition
  allows you to customize MobaXterm for your company: you can add
  your own logo, your parameters, your welcome message and generate
  either an MSI installation package or a portable executable.
  We can also modify MobaXterm or develop the plugins you need.
  For more information: https://mobaxterm.mobatek.net/download.html
```

/drives/c/Users/Elena van Zalen/OneDrive/Bureaublad
21/05/2025   13:25.20

Umeå Plant Science Centre

Software carpentry tutorial

Go to this link:

https://swcarpentry.github.io/shell-novice/index.html

Follow the instructions, completing these two sections:

1: Introducing the Shell

2: Navigating Files and Directories

Umeå Plant Science Centre

What is nf-core?

From their website: *A global community effort to collect a curated set of open-source analysis pipelines built using Nextflow.*

When starting a new project requiring bioinformatics analyses, it is not a bad idea to check the nf-core website to see if they have a pipeline prepared to run the analysis of the type of data you have:
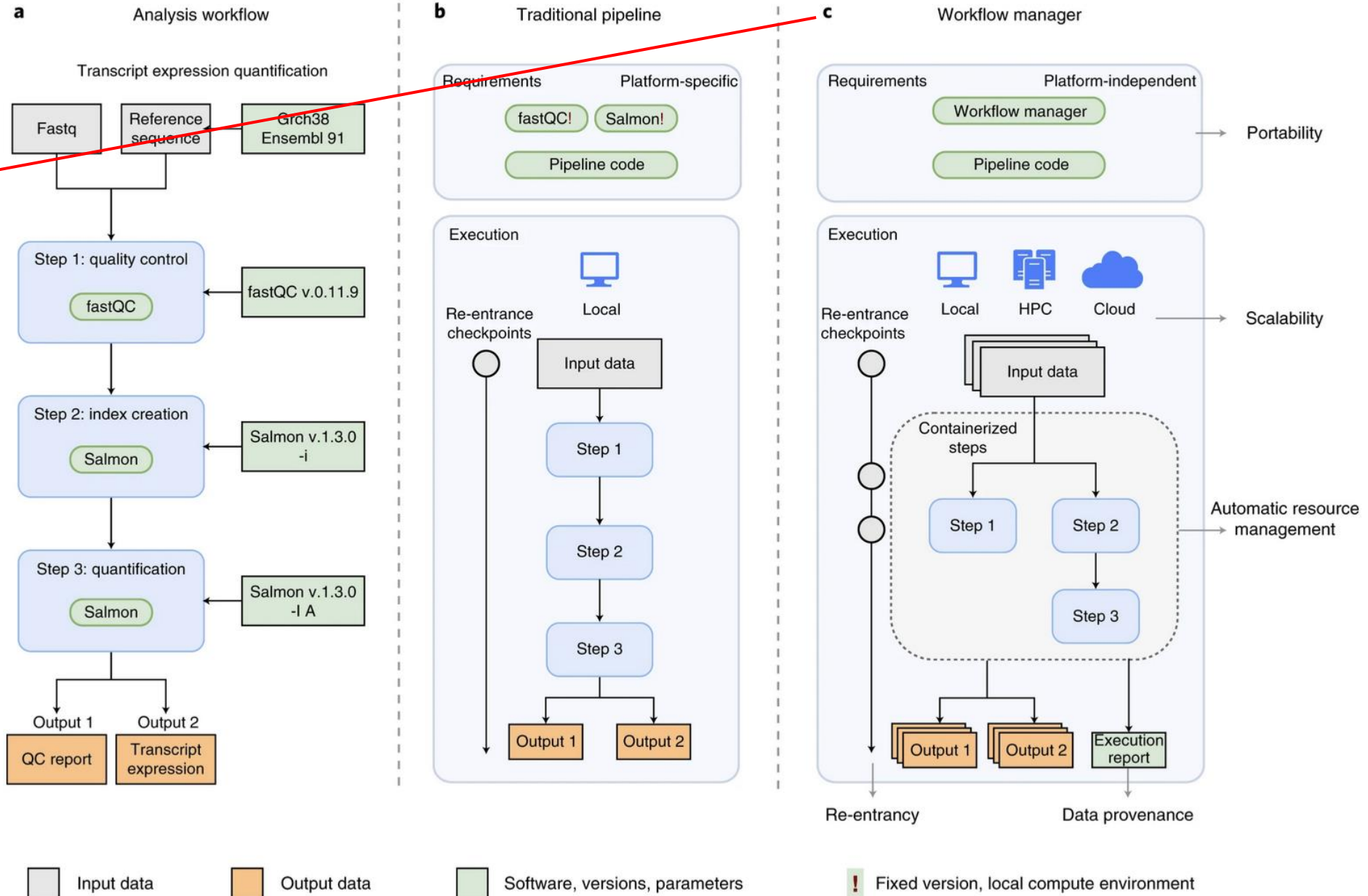
https://nf-co.re/

Umeå Plant
Science Centre

Visually:

Nextflow is the software allowing this

Nf-core is a repository of nextflow pipelines built for different bioinformatics applications

a Analysis workflow

Transcript expression quantification

Fastq — Reference sequence — Grch38 Ensembl 91

Step 1: quality control
fastQC ← fastQC v.0.11.9

Step 2: index creation
Salmon ← Salmon v.1.3.0 -i

Step 3: quantification
Salmon ← Salmon v.1.3.0 -I A

Output 1 — QC report
Output 2 — Transcript expression

b Traditional pipeline

Requirements — Platform-specific
fastQC! Salmon!
Pipeline code

Execution
Local
Re-entrance checkpoints — Input data
Step 1
Step 2
Step 3
Output 1 — Output 2

c Workflow manager

Requirements — Platform-independent
Workflow manager
Pipeline code → Portability

Execution
Local HPC Cloud → Scalability
Re-entrance checkpoints — Input data
Containerized steps
Step 1 — Step 2
Step 3 → Automatic resource management
Output 1 — Output 2 — Execution report
Re-entrancy — Data provenance

Input data Output data Software, versions, parameters ! Fixed version, local compute environment

Umeå Plant Science Centre

How to edit files

- In the rest of the lesson, you will edit a variety of files. They are all txt files organized in specific ways, such as **json** files.

- To edit them, use simple text editors and not word, as word tends to add metacharacters that are not correctly interpreted by the softwares we will use.

- Good text editors to use are TextEdit (for Mac) and Notepad or WordPad (for windows).

- A more complex one to use (very good, but with many functionalities that may be confusing) is VScode (https://code.visualstudio.com/download), which works on every system.

Umeå Plant
Science Centre

Running the rnaseq pipeline

- Download my github repository by entering in your shell the following command:

git clone https://github.com/EdoardoPiombo/nf-core-introduction.git

- Follow the instructions in the READme file to do the exercise.

- You have solutions in the "solutions" folder if you are stuck.

- When you are done, feel free to check out the other pipelines and see if there is something that could help you with your work:

https://nf-co.re/pipelines/

Umeå Plant
Science Centre