

Introducing the terminal

Data plumbers' corner, session 2

Arianna Masciolini

Språkbanken Text, University of Gothenburg

Hello world!

- ▶ Arianna Masciolini
- ▶ background in Computer Science
- ▶ PhD student in Language Technology at Språkbanken Text
(University of Gothenburg, Sweden)
- ▶ main interests: Computational Syntax and SLA
- ▶ met Ludovica at the UniDive training school in Chișinău
last summer

Today's session

1. What is the terminal and why do we care?
2. Basic bash commands and how to combine them
3. Text processing example

What you will need

A UNIX-like shell (preferably **bash**) or, if you don't have one installed on your laptop, [an online one](#):



What is a terminal?



What is a terminal?

Today, a terminal *emulator* is:

- ▶ an application that runs on your computer like any other application
- ▶ depending on what operating system you are using, it might be called: *terminal*, *command line/prompt* or *shell*¹
- ▶ a textual interface to your own machine

¹ but technically, shell ≠ terminal

Why do we still use a terminal?

- ▶ many useful programs only have a textual interface
- ▶ certain operations are more efficient (or just easier!) if performed through the command line
- ▶ it is the best way to run your own programs
- ▶ you will often interact with servers via the command line

What about bash?

- ▶ bash is a very popular *command-line shell* and *scripting language*
- ▶ it was an important part of **the UNIX operating system**
- ▶ today, it is preinstalled on most Linux (→ basically every server in the galaxy) and MacOS systems, but not on Windows (hence WSL)

Alice in Unixland

```
cat alice.txt | tr " " "\n" | sort  
| uniq -c | sort -nr | head -n10
```

Exercises

Modify the text processing command so to:

- ▶ show the 20 most frequent words
- ▶ show the 20 *least* frequent words
- ▶ show all words and their counts *in alphabetical order*
- ▶ write the results to a new text file
- ▶ show the 10 most frequent words, *excluding punctuation*
- ▶ show all words *whose frequency is exactly 1*
- ▶ ...

Next time

Version control with git, so:

- ▶ ensure you have bash (or very similar) installed
- ▶ save your scripts from today!