

The ABC of Computational Text Analysis




#3 Setting up your Development Environment

Alex Flückiger

Faculty of Humanities and Social Sciences
University of Lucerne

13 March 2025

Outline

- recap + discuss paper 
- interacting with computers 
- installation of programs 

Recap last lecture

- textual data is challenging 🤯
- machine-readable data 🚧
- reproducible research using computational approach ♻️

Reading

Computational Social Science (Lazer et al. 2009)

- data-driven
- network analysis + text analysis
- historical perspective vs. real-time dynamics
- issues: limited access to data and new methods

Computers



Two trends in computing

- everyday, mobile, cloud-centered, touch/click-based approach
 - single-purpose application
 - smartphones, smart assistants, augmented reality
- engineering, open-source, machine-oriented approach
 - powerful due to modularity
 - workflow for data science

(see also Healy 2019)

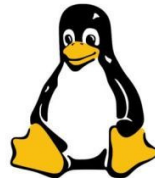


*Apple Original iPhone
(2007)*

Operating Systems (OS)

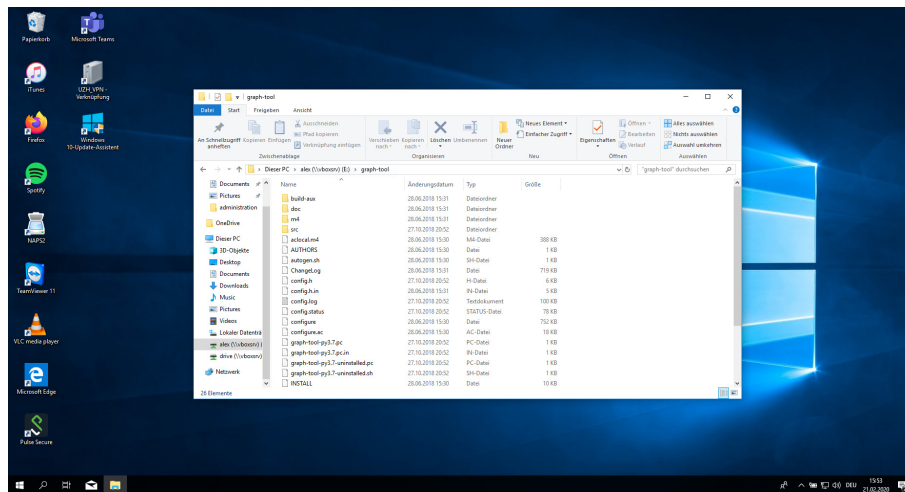
Between hardware and applications

- Microsoft Windows
- macOS
- Linux-based systems 🧐



User Interfaces (UI)

Graphical User Interface (GUI) Command Line Interface (CLI)



```
File Edit View Bookmarks Plugins Settings Help
New Tab Split View
Copy Paste Find
```

```
[alex@alex-laptop ~]$ cd KED2023
[alex@alex-laptop KED2023]$ ls -lh
total 340K
drwxr-xr-x 3 alex alex 4.0K Nov  4 18:46 bulk
drwxr-xr-x 7 alex alex 4.0K Mar 16 10:27 docs
-rw-r--r-- 1 alex alex 6.1K Feb 17 15:21 dodo.py
drwxr-xr-x 8 alex alex 4.0K Mar 16 10:28 KED2023
-rw-r--r-- 1 alex alex 1.3K Feb 23 14:05 lecture_intro.md
drwxr-xr-x 2 alex alex 4.0K Feb 17 15:21 lib
-rw-r--r-- 1 alex alex 6.9K Oct 21 15:57 LICENSE
-rw-r--r-- 1 alex alex 288K Feb 17 17:00 poetry.lock
drwxr-xr-x 2 alex alex 4.0K Feb 17 16:05 __pycache__
-rw-r--r-- 1 alex alex 922 Feb 17 17:02 pyproject.toml
-rw-r--r-- 1 alex alex 2.0K Feb 17 15:21 README.md
-rw-r--r-- 1 alex alex 345 Feb 17 15:21 requirements.txt
[alex@alex-laptop KED2023]$
```


«An awesome programmer saves the world»

Human-Machine Interaction

Tell the computer efficiently 🎯

- text commands instead of clicks
- programming similar to human languages
 - syntax (form)
 - semantics (meaning)
- computers are literalists

Programming

- **command languages**
Bourne-Again shell (Bash), since 1989
- **programming languages**
Python, R, Java, C++ etc.
- **instructions → program \equiv algorithm**
analogous to cooking

Package manager

App stores for engineers

- automation of software installation
 - system: apt, Homebrew
 - Python: pip, poetry, conda
- packages have dependencies

Open-Source is a mindset

Standing on the shoulders of giants

- free + open software
- cross-platform portability
- collaboration
 - share + reuse
- helpful community



Organizing data and code

Publishing code and data are key to open science. 🧐

- **Git** tracks file changes and supports version management
- **GitHub** is a popular hosting platform based on Git
 - share code and collaborate
 - repository = project folder

Resources everyone is using

- [Stack Overflow](#)
where you get answers to technical questions
- [GitHub](#)
where you find open software
- Chatbots (ChatGPT, Gemini, Deepseek etc.)
ask anything if you get stuck

Learning by doing, doing by Googleing +
ChatGPTing.



Set up your system

Backup

You don't need it until you desperately need it!

Setting up your development environment

Installation

- Python 3
 - packages for NLP + visualization
- various Bash tools
- VS Code Editor

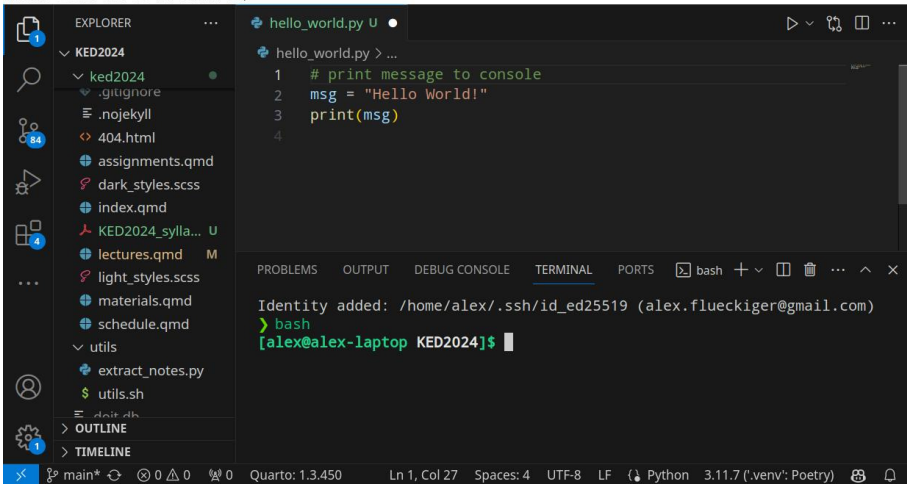
👉 Follow the [installation guide](#) for your OS + the Python chapter.

Relax.
It takes a while.

VS Code Editor

The Microsoft Word for coding

- powerful integrated development environment (IDE)
- cross-platform
- interactive analysis



The screenshot displays the Visual Studio Code (VS Code) editor interface. On the left, the Explorer sidebar shows a project named 'KED2024' with various files including '.gitignore', '.nojekyll', '404.html', 'assignments.qmd', 'dark_styles.scss', 'index.qmd', 'KED2024_sylla...', 'lectures.qmd', 'light_styles.scss', 'materials.qmd', 'schedule.qmd', 'utils', 'extract_notes.py', and 'utils.sh'. The main editor window is open to 'hello_world.py', which contains the following Python code:

```
1 # print message to console
2 msg = "Hello World!"
3 print(msg)
4
```

Below the editor, the TERMINAL panel is active, showing the output of a 'bash' command. The terminal text reads: 'Identity added: /home/alex/.ssh/id_ed25519 (alex.flueckiger@gmail.com)', followed by a prompt '[alex@alex-laptop KED2024]\$'.

Write your first Python script in [VS Code](#)

First steps in Python

Learn Python with interactive tutorials

- [Python Principles](#)
- [LearnPython](#)

Readings

Optional: pimp your workflow

Healy, Kieran. 2019. “The Plain Person’s Guide to Plain Text Social Science.”
[online.](#)



Questions?

References

Healy, Kieran. 2019. “The Plain Person’s Guide to Plain Text Social Science.”

<https://kieranhealy.org/publications/plain-person-text/>.

Lazer, David, Alex Pentland, Lada Adamic, Sinan Aral, Albert-László Barabási, Devon Brewer, Nicholas

Christakis, et al. 2009. “Computational Social Science.” *Science* 323 (5915): 721–23.

<https://doi.org/10.1126/science.1167742>.