

KNOWDIVE



KDI ● **Knowledge and Data Integration**

Data Management Tools & Libraries

iTelos Inception Phase

Alessio Zamboni

Contents

- 1 General environment**
- 2 Tools' Demo**
- 3 Libraries**
- 4 Libraries' usage Demo**

Language and tools

In order to handle data in the process we will use:

- Python 3.x [Language]
- Jupyter Notebook [IDE]
- Poetry [Dependency Manager]
- Git(Hub) [Versioning system]

Note

This is what we propose for this course because their simplicity, in real-world scenario those tools may vary depending on your requirements.

Python installation

Regardless your OS, there are mainly 2 ways for installing the programming environment:

- Conda (<https://bit.ly/2GpFg40>)
- Native package manager (depending on your OS)

Important Note

If you already have python3 installed on your system **please DO NOT MIX dependencies** from various sources, it works as follow:

- You downloaded Conda → Install packages with Conda
- You use the native installation → Install packages using Pip

Installation of the tools

After you have installed Python the basic tools has to be installed on top of the installation, to make it available for all projects.

- Jupyter Notebook (<https://jupyter.org/install.html>)
- Poetry (<https://bit.ly/3lj3zzv>)
- Git (not Python dependant, <https://bit.ly/3lclvww>)

Installation process is not covered here, the links are providing detailed information on how to proceed.

Live demo

Task specific libraries

Not exhaustive lists of suggested libraries:

Dataframe Pandas

Calcs NumPy

Plotting Matplotlib, graphviz

REST API Requests

Scraping Scrapy, BeautifulSoup4

Dates Dateparser, Arrow

Geospatial GeoPandas, geopy, geoplot (Cartopy)

Live demo



KDI : Knowledge and Data Integration



Alessio Zamboni



**Data Management Tools &
Libraries**

iTelos Inception Phase