

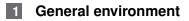
KDI **\*** Knowledge and Data Integration

# Data Management Tools & Libraries

iTelos Inception Phase

Alessio Zamboni

#### Contents



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 realworld 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:

- $\blacksquare$  You downloaded Conda  $\rightarrow$  Install packages with Conda
- $\blacksquare$  You use the native installation  $\rightarrow$  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/3lclvvw)

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

#### Live demo

Alessio Zamboni

## Task specific libraries

Not exhaustive lists of suggested libraries:

Dataframe Pandas

Calcs NumPy

Plotting Matplotlib, graphviz

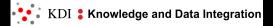
**REST API Requests** 

Scraping Scrapy, BeautifulSoup4

Dates Dateparser, Arrow

Geospatial GeoPandas, geopy, geoplot (Cartopy)

#### Live demo



# Alessio Zamboni

# Data Management Tools & Libraries iTelos Inception Phase