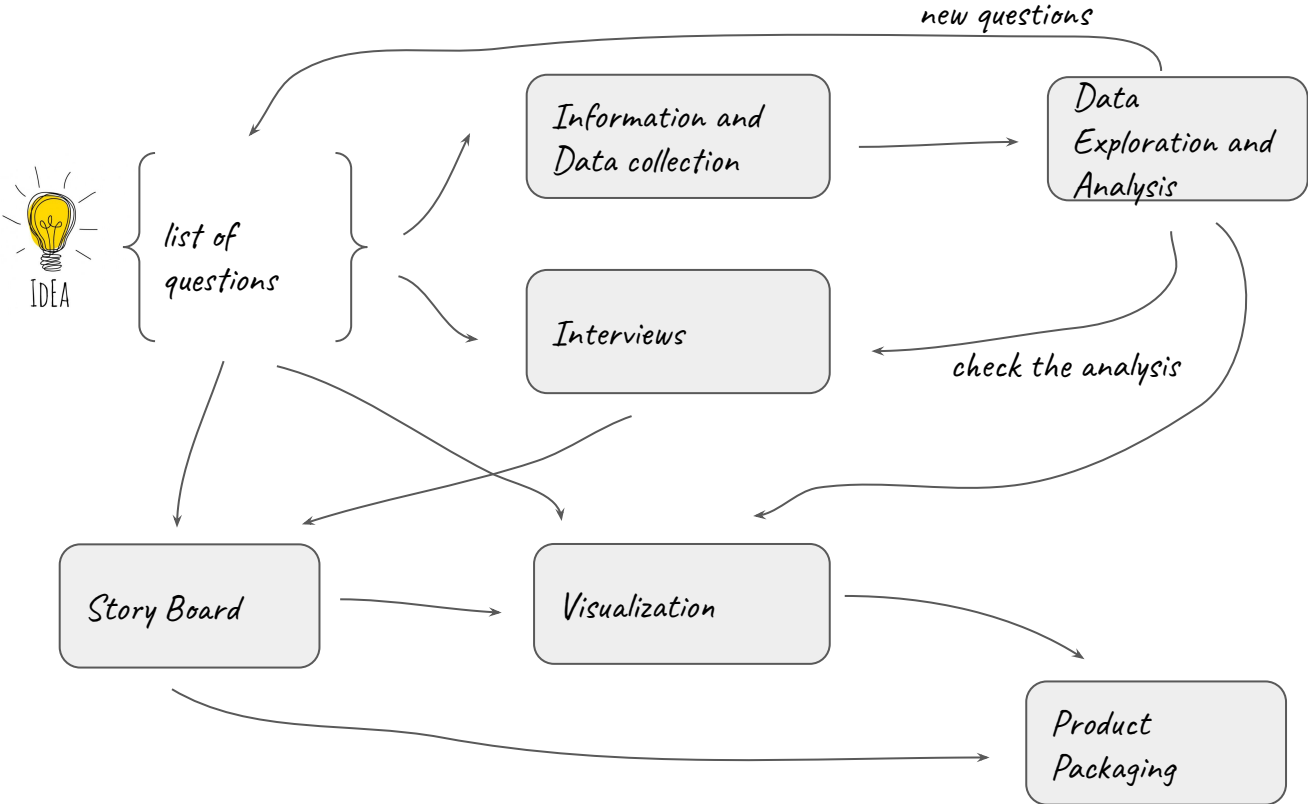


# Installazione del software del Data Journalist

Angelica Lo Duca  
angelica.loduca@iit.cnr.it

# Data Journalism Workflow



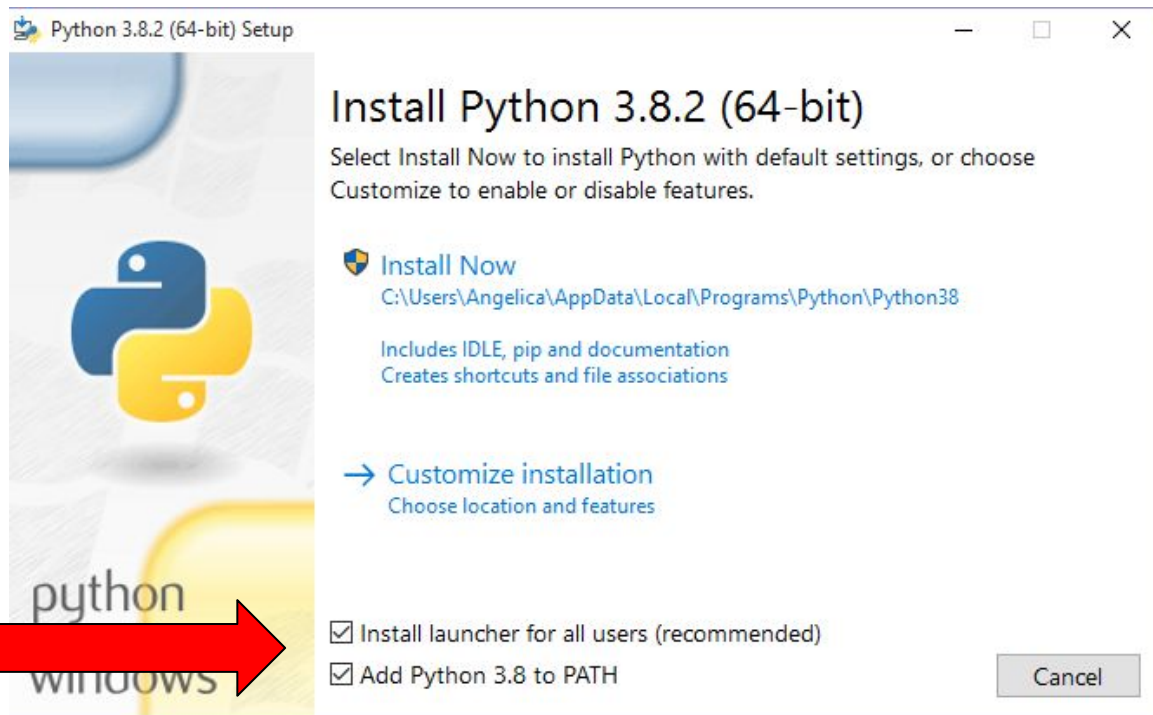
<b>Fase</b>	<b>Software</b>
Data Collection	python + jupyter + librerie specifiche
Data Cleaning	Open Refine
Data Exploration	Tableau
Data Analysis	python + jupyter + scikit-learn
Data Visualization	jupyter + D3.js
Story Packaging	XAMPP (opzionale), HTML + js

# Python 3 - Windows

- <https://www.python.org/downloads/> (pagina generale)
- <https://www.python.org/downloads/release/python-382/> (pagina specifica)

Version	Operating System	Description	MD5 Sum	File Size	GPG
<a href="#">Gzipped source tarball</a>	Source release		f9f3768f757e34b342dbc06b41cbc844	24007411	<a href="#">SIG</a>
<a href="#">XZ compressed source tarball</a>	Source release		e9d6ebc92183a177b8e8a58cad5b8d67	17869888	<a href="#">SIG</a>
<a href="#">macOS 64-bit installer</a>	Mac OS X	for OS X 10.9 and later	f12203128b5c639dc08e5a43a2812cc7	30023420	<a href="#">SIG</a>
<a href="#">Windows help file</a>	Windows		7506675dcbb9a1569b54e600ae66c9fb	8507261	<a href="#">SIG</a>
<a href="#">Windows x86-64 embeddable zip file</a>	Windows	for AMD64/EM64T/x64	1a98565285491c0ea65450e78afe6f8d	8017771	<a href="#">SIG</a>
<a href="#">Windows x86-64 executable installer</a>	Windows	for AMD64/EM64T/x64	b5df1cbb2bc152cd70c3da9151cb510b	27586384	<a href="#">SIG</a>
<a href="#">Windows x86-64 web-based installer</a>	Windows	for AMD64/EM64T/x64	2586cdad1a363d1a8abb5fc102b2d418	1363760	<a href="#">SIG</a>
<a href="#">Windows x86 embeddable zip file</a>	Windows		1b1f0f0c5ee8601f160cfad5b560e3a7	7147713	<a href="#">SIG</a>
<a href="#">Windows x86 executable installer</a>	Windows		6f0ba59c7dbeba7bb0ee21682fe39748	26481424	<a href="#">SIG</a>
<a href="#">Windows x86 web-based installer</a>	Windows		04d97979534f4bd33752c183fc4ce680	1325416	<a href="#">SIG</a>

# Installazione Python 3 su Windows



Selezionare entrambe le caselle con il segno di spunta.

# Verifica della correttezza dell'installazione

- Aprire un terminale
  - cercare nella barra di ricerca la parola “terminale”
  - selezionare prompt dei comandi
- Digitare **python --version**
  - come output dovrebbe dare: Python 3.8.2

# Python 3 - Mac OS

- <https://www.python.org/downloads/> (pagina generale)
- <https://www.python.org/downloads/release/python-382/> (pagina specifica)

Version	Operating System	Description	MD5 Sum	File Size	GPG
<a href="#">Gzipped source tarball</a>	Source release		f9f3768f757e34b342dbc06b41cbc844	24007411	<a href="#">SIG</a>
<a href="#">XZ compressed source tarball</a>	Source release		e9d6ebc92183a177b8e8a58cad5b8d67	17869888	<a href="#">SIG</a>
<a href="#">macOS 64-bit installer</a>	macOS	Python 3.8.2 and later	f12203128b5c639dc08e5a43a2812cc7	30023420	<a href="#">SIG</a>
<a href="#">Windows help file</a>	Windows		7506675dcbb9a1569b54e600ae66c9fb	8507261	<a href="#">SIG</a>
<a href="#">Windows x86-64 embeddable zip file</a>	Windows	for AMD64/EM64T/x64	1a98565285491c0ea65450e78afe6f8d	8017771	<a href="#">SIG</a>
<a href="#">Windows x86-64 executable installer</a>	Windows	for AMD64/EM64T/x64	b5df1cbb2bc152cd70c3da9151cb510b	27586384	<a href="#">SIG</a>
<a href="#">Windows x86-64 web-based installer</a>	Windows	for AMD64/EM64T/x64	2586cdad1a363d1a8abb5fc102b2d418	1363760	<a href="#">SIG</a>
<a href="#">Windows x86 embeddable zip file</a>	Windows		1b1f0f0c5ee8601f160cfad5b560e3a7	7147713	<a href="#">SIG</a>
<a href="#">Windows x86 executable installer</a>	Windows		6f0ba59c7dbeba7bb0ee21682fe39748	26481424	<a href="#">SIG</a>
<a href="#">Windows x86 web-based installer</a>	Windows		04d97979534f4bd33752c183fc4ce680	1325416	<a href="#">SIG</a>

# Verifica della correttezza dell'informazione

- Aprire un terminale
- Mac OS di default ha installato python 2.7
  - digitando **python --version** dovrebbe dare come output Python 2.7.10
- Per usare python 3 si deve usare il comando python3
  - digitando **python3 --version** dovrebbe dare come output Python 3.8.2



# Python 3 - Ubuntu

- Seguire la procedura disponibile a questo link:
  - <https://tecadmin.net/install-python-3-8-ubuntu/>

# Jupyter

- <https://jupyter.org/install>
  - pip install jupyterlab
  - pip install notebook
  - pip install voila

# Data Collection

	<b>Librerie</b>
Web API Flickr	pip install flickrapi
Web API Twitter	pip install python-twitter
Estrazione da PDF	pip install tabula-py tabula - <a href="https://tabula.technology/">https://tabula.technology/</a>
Web Scraping	selenium + chrome driver

# Selenium + Chrome Driver per python

- Da terminale digitale
  - **pip install selenium**
  - <https://selenium-python.readthedocs.io/installation.html>
- Web Driver per python
  - verificare la versione del browser

<b>Chrome:</b>	<a href="https://sites.google.com/a/chromium.org/chromedriver/downloads">https://sites.google.com/a/chromium.org/chromedriver/downloads</a>
<b>Edge:</b>	<a href="https://developer.microsoft.com/en-us/microsoft-edge/tools/webdriver/">https://developer.microsoft.com/en-us/microsoft-edge/tools/webdriver/</a>
<b>Firefox:</b>	<a href="https://github.com/mozilla/geckodriver/releases">https://github.com/mozilla/geckodriver/releases</a>
<b>Safari:</b>	<a href="https://webkit.org/blog/6900/webdriver-support-in-safari-10/">https://webkit.org/blog/6900/webdriver-support-in-safari-10/</a>

# Versione di Chrome

- Per conoscere la versione di Chrome, nella barra degli indirizzi digitare
  - `chrome://settings/help`



## Google Chrome

Aggiornamento quasi terminato. Riavvia Google Chrome per completarlo. Le finestre di navigazione in incognito non verranno riaperte.



Versione 80.0.3987.116 (Build ufficiale) (a 64 bit)

Riavvia

# Installazione Web Driver

- Dopo aver scaricato il Web driver, posizionarlo in una cartella a piacere del proprio file system
- Impostare la variabile di ambiente \$PATH in modo da contenere il percorso alla cartella prescelta
  - Windows -
  - Mac OS - modificare il file .profile
    - aprire un terminale
    - nella propria home digitare il comando **vi .profile**
    - in fondo al file scrivere
      - digitare il tasto i (insert)
      - export PATH="`<percorso al web driver>:$PATH`"
      - premere ESC :wq per salvare il file
    - riavviare il computer
    - aprire il terminale e digitare echo \$PATH

# Data Cleaning

- Open Refine
  - <https://openrefine.org/download.html>

# Data Exploration

- Tableau
  - <https://public.tableau.com/en-us/s/download>



# Data Analysis

- Scikit-learn
  - <https://scikit-learn.org/stable/install.html>
  - **pip install -U scikit-learn**

# Story Packaging

- XAMPP
  - <https://www.apachefriends.org/it/download.html>

# Conflitti con altri programmi

- Di default il server Apache sta in ascolto sulla porta 80.
- La porta 80 potrebbe essere utilizzata da altri programmi, come ad esempio Skype e Spotify.
- Per evitare conflitti, chiudere Spotify e i programmi che stanno in ascolto sulla porta 80.

# Conflitto con Skype

- Eseguire Skype
- Accedere al menu Strumenti e poi Opzioni
- Selezionare menu Avanzate e poi Connessione
- Deselezionare la voce evidenziata in ROSSO (**Usa le porte 80 e 443 per le connessioni in ingresso aggiuntive**)

