

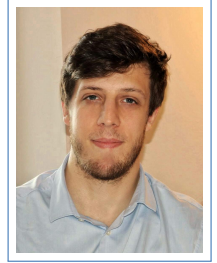
# Silvio Pavanetto

## Curriculum Vitae

6th May 1994

✉ [silvio.pavanetto@gmail.com](mailto:silvio.pavanetto@gmail.com)

🌐 [website / portfolio](#)



### Education

- 2016–2019 **Master of Science degree**, *Computer Science and Engineering, Politecnico di Milano, Italy. Grade: 110/110 with honors.*
- 2013–2016 **Bachelor of Science degree**, *Engineering of Computing Systems, Politecnico di Milano, Italy.*

### Work Experience

- November 2021 - Today **Software Engineer**, *Jobtome, Stabio, Switzerland.*
- Developed and deployed, on **GCP**, a job openings categorization model, using NLP deep learning techniques, that classifies approximately **1 million of jobs per day**.
  - Developed a scalable software architecture composed by more than **400 crawlers** for many different types of web pages, scheduled and executed in production using **Docker, Airflow and Kubernetes**.
- October 2019 **Research Fellow**, *Politecnico di Milano, Milan, Italy.*
- Worked on European projects mainly focused on data ingestion, analysis, and visualization, in collaboration with other universities.
  - Cooperated with the rest of my team on every aspect of the projects: requirements, system architecture design, development of software components.
  - Main technologies and libraries used: **Python, Flask, FastAPI, Docker, Pandas, Numpy, Matplotlib, Scikit-Learn, Tensorflow, Keras**.
- February 2019 **Software Developer**, *Deloitte Digital, Milan, Italy.*
- August 2019 **Software Developer**, *Deloitte Digital, Milan, Italy.*
- Worked as a developer on international cloud projects using **Salesforce** technology, which includes several programming languages such as **Java** and **Javascript**.

### Publications

- ICWE 2020 - Helsinki, Finland **Generation of Realistic Navigation Paths for Web Site Testing using Recurrent Neural Networks and Generative Adversarial Neural Networks.**
- Generated high-quality weblog data using deep learning techniques. Compared the results with a suite of data mining algorithms as a baseline. Main types of algorithms used: Recurrent Neural Networks and Generative Adversarial Neural Networks with a reinforcement learning approach. **Link**

- CySoc 2021 - *VaccinItaly: monitoring Italian conversations around vaccines on Twitter and Facebook*;  
 Online Monitored 3 millions of tweets and 1 million of Facebook posts of Italian users around vaccines in order to understand the interplay between the public discourse on online social media. **Link**
- ICWSM 2021 *A Content-based Approach for the Analysis and Classification of Vaccine-related Stances on Twitter: the Italian Scenario*;  
 - Online Collected and analyzed Italian 3 millions of conversations about COVID-19 vaccines on Twitter, investigating the geographical, temporal and lexical distribution of data. Trained a binary classifier that predicts the stance of tweets towards vaccines, i.e., it applies a "Pro-vax" or "No-vax" label. **Link**
- CSCW 2021 - *The Contribution of Textual Data from Online Reviews for the Evaluation of Service Quality: The Experience of Italian Museums*;  
 Online Investigated service quality dimensions of museums through a content analysis based on supervised and unsupervised models for the text of online reviews of 100 Italian museums over a time-period of one year. **Link**

## Projects

- *News and Social Media Data Analysis Pipeline related to COVID-19*: development of data collection and analysis pipeline for studying COVID-19 related contents and web dashboard available for citizens. **NLP** techniques and relative Python implementations (**NLTK**, **BERT**, **word2vec**) involved in the project. Tools used for building microservices architecture: **Docker**.  
**Github Link**
- *Italian Museums Reputation*: Development of data collection pipelines and periodical data ingestion systems using **MongoDB**, **Python**, and web scraping techniques (**BeautifulSoup**, **Selenium**, **Requests**).
- *VaccinItaly*: Monitor Italian conversations around vaccines on social media (Twitter, Facebook) and understand the interplay between online public discourse and vaccine hesitancy/uptake rates. Techniques and tools used: **MongoDB**, **Python**, **Scikit-Learn**, **Pandas**, **Transformers**, **Keras**, **BERT**.  
**Github Link**
- *Fanta NBA Predictor*: Prediction of fantasy scores of NBA players in the real games, using data science and machine learning techniques. Collected a huge amount of statistics from different sources and then used them to perform the analysis and the predictions. Main languages and frameworks used: **Python**, **Numpy**, **Pandas**, **Scikit-Learn**, **BeautifulSoup**, **Keras**.  
**Github Link**

## Hobbies and Passions

Basketball (I've been playing It since I was 8 years old), Chess, Logical Problems / Puzzles / Riddles, Music (of all kinds): I've played violin and percussion for a few years.

## Languages

Italian **Native Speaker**  
 English **Proficient**