

Flavia Salutari

PHD IN COMPUTER SCIENCE | DOB: 30/11/1993

☎ (+39) 3297088802 - (+33) 0668954253 | ✉ flavia.salutari@gmail.com | 🏠 flviasalutari.github.io/ | 📱 [flviasalutari](#) | 🎓 [flviasalutari](#)

PhD in computer science, data and AI at Institut Polytechnique de Paris. Focused for the past 4 years on natural language processing for Web-based content and on mining of large scale Internet measurements (host classification, users' satisfaction prediction) via machine learning. Proudly published in conferences and journals I admire like The Web Conference.

Professional Experience

Institut Polytechnique de Paris (Telecom Paris)

Paris, France

PH.D. CANDIDATE IN COMPUTER SCIENCE, DATA AND AI

Jun. 2018 - Jun. 2021

- Defined methodologies and tools for a large scale, longitudinal and unbiased view of Internet (protocols and user experience).
- Led experiments to capture the Quality of Experience (**QoE**) of users browsing the Web, with both online and offline experiments.
- Built supervised data-driven models (**Random Forest, XGBoost, k-NN, SVM, MLP**) to forecast users' satisfaction based on objective metrics in the context of a 2 years long collaboration with the [Wikimedia Foundation](#) to explore and predict Wikipedia users' QoE.
- Published *Analyzing Wikipedia Users' Perceived Quality Of Experience: A Large-Scale Study* for IEEE Transactions on Network and Service Management (2020), nominated for the IEEE Communications Society William R. Bennett Prize.
Presented *A large-scale study of Wikipedia users' quality of experience* at The Web Conference (WWW), San Francisco, CA, 2019.
Presented *Implications of the Multi-Modality of User Perceived Page Load Time* at IEEE MedComNet Conference, Online, 2020.

University College London

London, UK

VISITING RESEARCHER

Sep. 2020 - Feb. 2021

- Investigated and revealed societal biases embedded in Natural Language Processing (**NLP**) algorithms for dialogue systems.
- Led experiments (masked language modeling and language generation) to measure and analyze the fairness of state-of-the-art transformer-based language models (**BERT, BART, GPT-2**, etc.) towards renowned datasets of spoken English (7M+ data points).

Telecom ParisTech

Paris, France

R&D ENGINEER & INTERN (RESEARCH FUNDINGS: NEWNET@PARIS CISCO'S CHAIR)

Mar. 2017 - May 2018

- Independently built a framework to run large scale Internet measurement campaigns via active probing for host classification.
- Exposed the first Internet-wide adoption (6M+ representative targets) of both known and undiscovered IP-ID implementations.
- Presented *A closer look at IP-ID behavior in the Wild* at the International PAM Conference (PAM), Berlin, Germany, 2019.

Telecom Italia

Torino, Italy

INTERN

Apr. 2015 - Jun. 2015

- Designed and developed an algorithm (in MATLAB) for the geometric verification of pictures, in the context of image recognition.

Education

Politecnico di Torino

Sep. 2015 - Oct. 2017

MASTER OF SCIENCE IN ICT FOR SMART SOCIETIES (FINAL GRADE: 110/110 SUMMA CUM LAUDE)

Torino, Italy

- Thesis Title: *Design and experimentation of host classification algorithms using active measurements in the Internet.*
- Coursework: Operational Research, Statistical Signal Processing, Automotive Infosystems, Smart Grids, ICT in transport systems, ICT in building design, ICT for Geomatics, ICT for Health, Programming for IoT applications, Mobile and Sensor Networks.
- As a junior researcher worked with a team to build a framework (**Python, Pandas, GeoPandas, MongoDB**) for the collection and analysis of car sharing data: *UMAP: Urban Mobility Analysis Platform to Harvest Car Sharing Data* for IEEE Smart City Innovations (2017).

Tongji University

Sep. 2013 - Jul. 2014

POLITONG PROJECT (EXCHANGE PROJECT)

Shanghai, China

- Coursework: Complex Analysis, Advanced Programming, Software Engineering, Chinese Language, Chinese History and Culture.

Politecnico di Torino

Sep. 2012 - Jul. 2015

BACHELOR DEGREE IN TELECOMMUNICATIONS ENGINEERING

Torino, Italy

- Coursework: Calculus, Signal Theory, Computer Networks, Image and Video Processing, Networking Laboratory.

Skills and Interests

IT Tools

Git, SVN, Apache Hive, Spark, Microsoft Office

Programming

Python (Jupyter Notebooks/PyCharm/Spyder), Bash, \LaTeX , SQL, MATLAB

Python Libraries

Scipy (Pandas, Numpy, Matplotlib, Stats), Scikit-Learn, Pytorch, Keras, NLTK, Transformers, Tensorflow, Pyspark

Languages

Italian (native), English (advanced), French (advanced), Chinese (beginner)

Soft Skills

Communication, Teamwork, Empathy, Adaptability

Other Interests

Acting, Cinema, Theater, Photography, Psychology

Additional Activities

Telecom Paris ~ University College London

Paris (France), London (UK)

TEACHING ASSISTANT

Sep. 2017 - Feb 2021

- Courses: Geospatial Programming (2020), Exploration de grands volumes de données (2020), Mining of Massive Datasets (2019,2020), From Complexity to intelligence (2020), Internet Applications (2017, 2018).

Hackathon

Wien (Austria), Porquerolles (France)

ATTENDEE AND WINNER

Jun. 2018

- Participated to TMA 2018 (development of an algorithm for fingerprinting the usage of Internet services) and RESCOM 2018 (temporal and spatial analysis of tweets) hackathons in team with two other PhD candidates and won both.

I hereby consent to the use of this information according to the local privacy laws (e.g. GDPR 679/16).