

Silvia Casola

Contact:

- scasola@fbk.eu
- +39 3277408493
- [silvia.casola3](#) (Skype)

Personal details:

- Place of birth: Mazara del Vallo (TP), Italy
- Date of birth: 08 July 1993
- Citizenship: Italian
- Spoken Languages: Italian (Native), English (Proficient), Spanish (Conversant)

Updated: December 15, 2022

Research Interests

Natural Language Processing, Deep Learning, Machine Learning.

Research Profile

I am a Ph.D candidate in Natural Language Processing at University of Padua, owning a grant from Fondazione Bruno Kessler (Trento); my advisor is Alberto Lavelli (FBK). I have just concluded a 6-month visiting period at the TALN group in UPF (Barcelona); I also continue collaborating with Huawei Research Ireland after a 6-month internship. I obtained a Laurea Magistrale (Master Degree) with honours in Computer Engineering (Data Science Curriculum) from Politecnico di Torino. My research interests lay in the Natural Language Processing area, with a focus on summarization and simplification.

Education

- **University of Padua, Italy** *October 2019 – present*
Ph.D. in Brain, Mind and Computer Science
 - **Politecnico di Torino, Italy** *2016 – 2018*
Laurea Magistrale (2 years M.S. equivalent) in Computer Engineering (Data Science curriculum).
Thesis: *Reducing waiting times and crowding in hospital emergency departments using Machine Learning*.
Supervisors: Prof. Ricard Gavaldà (UPC) and Prof. Silvia Chiusano (Politecnico di Torino)
Grade: *110/110 cum laude*
 - **Politecnico di Torino, Italy** *2013 - 2016*
Laurea (3 years B.S. equivalent) in Computer Engineering.
Grade: *102/110*
-

Internships and Visiting Periods

- **Universitat Pompeu Fabra, Barcelona, Spain** *March 2022 – August 2022*
Visiting Ph.D. Student
I visited the TALN group, under the supervision of Prof. Saggion. I collaborated with the group researchers and developed my interest in summarization and simplification.
 - **Huawei Research, Ireland (remote)** *July 2021 – December 2021*
NLP Ph.D. Intern
I developed a module to correct misspelled queries in the AppGallery, which has been deployed. Research wise, I am working on ensuring factuality and minimizing hallucination in a sequence to sequence system for text summarization. I continue collaborating on summarization and factuality research.
 - **Universitat Politècnica de Catalunya, Spain** *September 2017 - July 2018*
Erasmus Program Master's student – MS in Artificial Intelligence, MS in Innovation and Research
I followed courses in the Machine Learning area and wrote and discussed my Master's thesis.
-

Teaching

- **Programming Teaching Assistant (Python)** *October 2020 - February 2021*
I gave classes, and assisted with exercises and homework grading (Programming course; first-year Math undergrads at University of Padua)

- **Computer Science Course Teaching Assistant (C)** *September 2017 - February 2018*
I assisted with exercises and other lab work (Computer Science course; first-year Engineering undergrads at Politecnico di Torino)
-

Supervision

- **Master's Student (internal supervisor)**
- Lucia Larocca, 2019 (Politecnico di Torino Master thesis)
-

Invited talks

- **Natural Language Processing: an overview**
Machine Learning Course, University of Padua (invited by Prof. F. Aioli)
-

Academic Appointments & Service

- **Research Assistant** *January 2019 - October 2019*
I was part of a team in Politecnico di Torino designing a recommendation system for health structures (CANP project). I received a scholarship / stipend for the whole period.
 - **Reviewer**
ESANN 2020, SMM4H 2020, GEM 2022
 - **PhD course student representative** *February 2022 - present*
-

Attended schools and courses

- **Technische Universität München (TUM), Germany** *August 2021*
1st Munich Legal Tech Summer School
 - **University of Grenoble and Naver Labs Europe, Online** *January 2021*
Advanced Language Processing winter School (ALPS)
 - **coursera.org, deeplearning.ai**
Deep Learning Specialization; held by Prof. Andrew Ng (Stanford University) - Certificate
Machine Learning Course; held by Prof. Andrew Ng (Stanford University)
-

Other activities

CD: 50/50, co-founder
No-profit association for promoting gender equality and diversity in STEM.
Organized and instructed short courses on Python and data science for high schoolers (Rome, 2022, 20h).

Skills

Programming languages: Python (first language), SQL, C. Basics in R, C++ and Java.
Machine/Deep Learning stack: Pytorch; Numpy, Pandas, Matplotlib, Scipy, Scikit-learn, etc.; Hugging Face ecosystem, NLTK, Spacy; Weight and Biases. Experience with large-scale GPU clusters.

Languages

Italian (native), English (full professional proficiency), Spanish (conversational proficiency)

List of Publications

- **Casola, Lavelli, Saggion (2022):** What's in a (dataset's) name? The case of BigPatent. Proceedings of the 2nd Workshop on Generation, Evaluation & Metrics (GEM), EMNLP 2022
- **Obonyo, Casola, Saggion (2022):** Exploring the limits of a base BART for multi-document summarization in the medical domain. Proceedings of the 3rd Workshop on Scholarly Document Processing, COLING 2022
- **Casola, Lavelli (2022):** Summarization, Simplification, and Generation: The Case of Patents. Expert Systems with Applications
- **Casola, Lauriola, Lavelli (2022):** Pre-Trained Transformers: An Empirical Comparison. Machine Learning with Applications

- **Casola**, Lavelli (2021): WITS: Wikipedia for Italian Text Summarization. Eighth Italian Conference on Computational Linguistics. CLiC-IT 2021
- Louvan, **Casola**, Magnini (2021): Investigating Continued pretraining for Zero-Shot Cross-Lingual Spoken Language Understanding. Eighth Italian Conference on Computational Linguistics. CLiC-IT 2021
- **Casola**, Lavelli (2020): FBK@SMM4H2020: RoBERTa for detecting medications on Twitter. Proceedings of the Fifth Social Media Mining for Health Applications Workshop & Shared Task, COLING 2020
- Bassi, **Casola**, Mancinelli, Lai, Salcuni (2020): Internalizing-Externalizing Symptoms as Predictors of Problematic Smartphone Use among Adolescents: A Machine Learning approach. International Congress of Clinical and Health Psychology in Children and Adolescents. Poster Presentation - Competing as award for young researcher
- Bazzano, Montuschi, Lamberti, Paravati, **Casola**, Ceròn, Londoño, Tanese (2017): Mental Workload Assessment for UAV Traffic Control Using Consumer-Grade BCI Equipment. In: Intelligent Human Computer Interaction. IHCI 2017.