

Paolo Gajo

☎ +1 902 448 5839 | @ paolo.gajo2@unibo.it |  LinkedIn |  GitHub | 📍 Forlì, Italy/Halifax, Canada

PhD student in NLP at the University of Bologna researching gastronomy through LLMs. Currently visiting the Hypermatrix group at Dalhousie University, Canada, researching graph neural networks for text-to-graph models. Personal research involves ASR and seq2seq models for automatic subtitling for the deaf and hard of hearing.

PUBLICATIONS

- Paolo Gajo and Alberto Barrón-Cedeño. On Cross-Language Entity Label Projection and Recognition. In *CLiC-it 2024: 10th Italian Conference on Computational Linguistics*, 2024.
- Paolo Gajo, Luca Giordano, and Alberto Barron-Cedeño. UniBO at CheckThat! 2024: Multi-lingual and multi-label persuasion technique detection in news with data augmentation and sequence-token classifiers. In Guglielmo Faggioli, Nicola Ferro, Petra Galuščáková, and Alba García Seco de Herrera, editors, *Working Notes of CLEF 2024 - Conference and Labs of the Evaluation Forum*, CLEF 2024, pages 426–434, Grenoble, France, 2024.
- Paolo Gajo, Arianna Muti, Katerina Korre, Silvia Bernardini, and Alberto Barrón-Cedeño. On the identification and forecasting of hate speech in incelism. In Ruslan Mitkov and Galia Angelova, editors, *Proceedings of the 14th International Conference on Recent Advances in Natural Language Processing*, pages 373–384, Varna, Bulgaria, September 2023. INCOMA Ltd., Shoumen, Bulgaria.
- Paolo Gajo, Silvia Bernardini, Adriano Ferraresi, Barrón-Cedeño Alberto, et al. Hate speech detection in an italian incel forum using bilingual data for pre-training and fine-tuning. In Federico Boschetti and Nicole Novielli Gianluca E. Leboni, Bernardo Magnini, editors, *CLiC-it 2023: 9th Italian Conference on Computational Linguistics*, volume 3596, Venice, Italy, 2023. CEUR-WS, CEUR-WS.
- Aikaterini Korre, Paolo Gajo, and Alberto Barrón-Cedeño. Hate Speech According to the Law: An Analysis for Effective Detection. In *Proceedings of the 31st International Conference on Computational Linguistics (COLING 2025)*, 2025. To appear.

PROJECTS

Automatic Subtitle Segmenter

2023

- Fine-tuning a T5 model on a dataset built from my work as a subtitler to build an automatic subtitle segmenter for audiovisual content. Code: [GitHub](#)

SKILLS

Programming	Python, R, Bash
Libraries	PyTorch, NumPy, Transformers, Pandas, Sci-Kit Learn, SpaCy, Selenium, BeautifulSoup
Software	Git, VS Code, LaTeX, Overleaf, Slurm, Sketch Engine, AntConc, RegEx, Excel, Label Studio, Trados, MemoQ, Subtitle Edit, Aegisub
Text processing	RegEx, Sketch Engine, AntConc, BootCaT
Translation	Trados, MemoQ, MateCat, MultiTerm, Subtitle Edit, Aegisub

EDUCATION

Dalhousie University

Halifax, Canada

Visiting @ [Hypermatrix](#)

Sep 2024 – Sep 2025

- Researching graph neural networks and text-to-graph models.
- Courses attended: MATH/STAT 2060 - Intro Probability & Statistics, CSCI 3151 - Foundations of Machine Learning, CSCI 4158 - NLP with Deep Learning, CSCI 6516 - Deep Learning.

University of Bologna*PhD in Natural Language Processing*

Forlì, Italy

2023 – Present

- 3-Year PhD on the use of Natural Language Processing for the study of gastronomy.

University of Bologna*MA in Specialized Translation (110/110 Summa Cum Laude)*

Forlì, Italy

2021 – 2023

- 2-Year degree on translation technologies and translation in English, Spanish, and Italian.
- Thesis: *This Is My Cope: Identification and Forecasting of Hate Speech in Inceldom*. Code: [GitHub](#)

Ca' Foscari University*BA in Linguistic and Cultural Mediation (110/110 Summa Cum Laude)*

Treviso, Italy

2018 – 2021

- 3-Year degree focusing on translation and interpreting in English, Spanish, and Italian.
- Thesis: *Practical Application of Subtitling Guidelines - Intralingual and Interlingual Subtitling of the Short Film "Group B"*