# Giulio Cusenza

giuliocusenza@gmail.com +39 334 980 3678



#### Education

Oct 2021 -Oct 2024 Bachelor's in Computational Linguistics @ University of Tübingen

Completed the Bachelor's in Computational Linguistics at the University of Tübingen with an average grade of **1.25** (**A**) within standard program duration.

Jun 2021

High school degree with A grade

Graduated with **100/100** at the **classical high school** "Liceo Classico Umberto I" (key subjects: **Latin**, **Ancient Greek**, **Philosophy**) in Palermo, Italy.

## Experience

Jan 2024 -May 2024 Intern in Data & Technology @ MHP – A Porsche Company

Automating company processes by applying **Machine Learning** and **NLP** to **Robotic Process Automation** with **UiPath**.

Oct 2022 -Dec 2023

Programming teaching assistant @ University of Tübingen

**Tutored over 70 students** in the courses "Data structures and algorithms for computational linguistics" **I, II (Java)** and **III (python)**.

Apr 2023 -Sep 2023 Research assistant @ University of Tübingen

Worked on a project on the **automatic identification** of **language borrowing** and **contact** through NLP and bioinformatic methods (DFG Center for Advanced Studies "Words, Bones, Genes, Tools").

#### **Skills**

IT:

- Python (advanced), Java (intermediate).
- Machine learning projects with pytorch.
- Experience with **regex**, **context-free** grammars, and **finite state automata**.
- Experience with web development: HTML, CSS, JS, Firebase (example).
- Familiarity with bash, R, SQL, Prolog.
- Version control: **Git**.
- Media: Microsoft Office, LaTeX.

Languages:

- Spoken: Italian (native), English (C2), French (C1), Russian (B2), German (B2).
- Natively understood: Sicilian, Arbëresh.
- Written: Latin, Ancient Greek.
- Basic: Dutch, Spanish, Korean.

Other: philosophy literary analysis

Traits: curious creative detail-obsessed committed communicative

## **Publications**

• Giulio Cusenza and Çağrı Çöltekin. 2024. <u>NLP for Arbëresh: How an Endangered Language Learns to Write in the 21st Century</u>. In *Proceedings of the 3rd Annual Meeting of the Special Interest Group on Under-resourced Languages* @ *LREC-COLING 2024*, pages 252–256, Torino, Italia. ELRA and ICCL.

# **Projects**

**Arbor** Website with self-developed NLP tools (spellchecking, conjugation...) for the extremely

low-resourced Arbëresh language spoken in Southern Italy.

Arboi

**Question tagger model** Training of a question labelling model based on Hugging Face Transformers.

link

**FST spell checker** Probabilistic spell checker built on weighted finite state transducers (python

implementation).

GitHub link

**Kneser-Ney bigram** 

model

Python program to train a word-bigram model with Kneser-Ney smoothing for next-

word prediction and sentence generation.

GitHub link

## **Awards**

**Deutschlandstipendium** (scholarship)

"The Deutschlandstipendium supports talented and high-performing students. In addition to first-class grades, social commitment and personal achievements are also considered for the award."



<u>link</u>