

# Matteo Guida

00 (49) 176 3077 0855 | ✉ matteo.guidaa@gmail.com | 🔗 linkedin.com/in/matteo-guida-7872341b3/ | 🐙 GitHub

## Personal Profile

---

Masters Student Graduate in Speech and Language Processing at the University of Konstanz (Germany) with a solid theoretical background in (computational) linguistics and knowledge in Deep Learning and statistical methods, and practical industry experience in the field of Natural Language Processing.

## Education

---

### University of Konstanz

Konstanz, Germany

Masters in Speech and Language Processing

Oct 2021 - Current

- Main track: Natural Language Processing
- **Relevant Courses:** Corpus Linguistics with Python, Deep Learning with Python, Machine Language Processing, Statistics with R, Linguistics, Finite State Morphology
- Average GPA: 1.2 (1 being the highest)
- Expected date of graduation: February 2024
- Thesis: Automatic Hate Speech Detection in Italian Text using Deep Learning Techniques (BERT and Llama)

### University of Trieste

Trieste, Italy

Bachelor in Applied Linguistics

Oct 2018 - July 2021

- Graduated with 108 out of 110
- Thesis: "German in a non-European multilingual context: the case of Namibia".

### Liceo Galileo Galilei

Civitavecchia, Italy

Italian-French High School

Sept 2013 - July 2018

- Graduated with 100 out of 100
- Received both the Italian and French high school diploma. Classes were taught in Italian, French, Spanish and English.

## Work Experience

---

### DHL Group

Bonn, Germany

Working Student, Natural Language Processing

September 2023 - January 2024

- Fine-tuning of LLMs for domain specific content generation, entity recognition and text summarization, Generative Question-Answering with LangChain, Custom Machine Translation & Evaluation, Data pre-processing, cleaning and annotation
- **Technical Skills:** Machine Learning, Kubernetes (Kubeflow), Azure, Large Language Models, Langchain, Generative AI, Deep Learning for NLP, Git, RestAPIs, AWS

### Deutsche Post DHL Group

Bonn, Germany

Intern in Linguistic Data Management

March 2023 - August 2023

- Data Cleaning & Annotation, Data-Specific Customized NER Models for terminology extraction, evaluation of models, (Neural) Machine Translation
- **Technical Skills:** Programming in Python, NLP libraries, Deep Learning, Git, RestAPIs, AWS, Google Cloud AutoML, Prodigy, Tableau

### University of Konstanz

Konstanz, Germany

Teaching Assistant for the MA Course "Machine Language Processing"

Oct 2022 - Feb 2023

- This course offers both a theoretical and practical introduction into Natural Language Processing and covers a variety of fundamental topics in the field of Speech and Language Processing (including Regular Expressions, Text Pre-Processing, supervised and unsupervised ML, information extraction, dialogue systems) and a practical introduction to Linux.
- **My tasks:** Weekly tutorial to help students with Linux-related issues (useful command lines and RegEx for computational linguists) and theoretical NLP concepts, correction and grading of assignments.

### University of Konstanz

Konstanz, Germany

Student Research Assistant

Nov 2022 - Feb 2023

- Project on sarcasm detection which aims at finding what contexts might trigger (or not) sarcasm and if multimodal clues also play a role in it.
- **Technical Skills:** Programming in Python and R (Data Cleaning & Preparation), Excel, evaluation of NLP models, Deep Learning (Word Embeddings, BERT, Transformers)

## Skills & Knowledge

---

<b>Programming</b>	Python (Pandas, PyTorch, NumPy, Scikit-learn, Spacy, NLTK, Core NLP, Stanza, Gensim, TensorFlow, Keras), R
<b>Miscellaneous</b>	Linux, Shell (Bash/Zsh), $\LaTeX$ (Overleaf/R Markdown), Microsoft Office, Git
<b>Machine Learning</b>	Neural Networks, Linear Regression, Logistic Regression, TF-IDF, evaluation metrics, Word Embeddings (Word2Vec, GloVe)
<b>&amp; Others</b>	Deep Learning architectures (CNNs, RNNs, LSTMs, Transformers), Prodigy, Tableau, APIs

## Workshops

---

### Basque Summer Tutorial in Language Documentation

*Larrau/Tardets-Sorholus, France*

*16 - 26 Aug 2022*

- Summer School in collaboration with the University of Bordeaux, the French-German University, the University of Konstanz, IKER, the University of Pau and Pays de l'Adour, and the CNRS.
- Conceptual, technical and ethical issues in Language Documentation and endangered languages, along with actual fieldwork in-situ (conducting interviews and analyzing data)
- Research Proposal: an analysis of vowels production in Zuberoan Basque
- **Technical Skills:** Praat, ELAN, Audacity, R

## Languages

---

<b>Italian</b>	Native proficiency
<b>English</b>	Full professional proficiency
<b>French</b>	Full professional proficiency
<b>German</b>	Full professional proficiency
<b>Spanish</b>	Full professional proficiency

## Interests

---

**Linux, Linguistics, Natural Language Processing, Ethical AI**

**References available upon request.**