Matteo Guida

📱 00 (49) 176 3077 0855 | ✉ matteo.guidaa@gmail.com | 🏠 https://cis.unimelb.edu.au/people/students/matteo-guida

Personal Profile

PhD candidate at the University of Melbourne, specializing in Computational Linguistics and Natural Language Processing within the Faculty of Engineering and Information Systems. My research focuses on developing automatic methods for detecting framing in news coverage of polarizing contemporary issues. I hold a Masters in Speech and Language Processing and have practical experience in Generative AI from my tenure at DHL, where I applied NLP and ML techniques to deliver practical solutions. My interests span polarization, misinformation, disinformation, propaganda, and societal dynamics, leveraging NLP to better understand human behavior.

Education

University of Melbourne
Melbourne, Australia

Ph.D. — *April 2024 - October 2027*

- Research: Towards Multilingual Models of Framing Detection in News Coverage of Contemporary Polarizing Issues
- School of Computing and Information Systems, NLP/AI group
- Supervisors: Prof. Lea Frermann, Prof. Eduard Hovy, Dr. Yulia Otmakhova

University of Konstanz
Konstanz, Germany

Masters in Speech and Language Processing — *Oct 2021 - February 2024*

- Main track: Natural Language Processing
- **Relevant Courses:** Corpus Linguistics with Python, Deep Learning with Python, Machine Language Processing, Statistics with R, Ethical AI, Linguistics, Finite State Morphology
- 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*

- Thesis: "German in a non-European multilingual context: the case of Namibia".

Work Experience

University of Melbourne
Melbourne, Australia

Research Assistant — *August 2024 - Present*

- Contributing to the project "Measuring Social Media Speed and the Acceleration of Informational Crisis," supervised by Dr. Philipp Pond.
- Conducting data collection, preprocessing, and analysis of social media platforms to study information dissemination and crisis acceleration.
- **Technical Skills:** Natural Language Processing, Social Media Analytics, Python, Data Cleaning and Annotation, Statistical Analysis, Data Visualization.

DHL Group
Bonn, Germany

Working Student, Natural Language Processing — *September 2023 - February 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, Kubernets (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.
- **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)

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, Audactiy, R

Technical Skills

Programming Languages	Python (Advanced), R (Intermediate), SQL, Shell Scripting
ML/AI Frameworks	PyTorch, TensorFlow, Transformers, LangChain, Kubeflow, Hugging Face
Cloud & DevOps	AWS, Azure, Google Cloud AutoML, Kubernetes, REST APIs
NLP	BERT, Llama, Word Embeddings, NER, Neural Machine Translation, LLMs
Languages	Italian (Native); English, French, German, Spanish (Full Professional)

References available upon request.