Gloria Galasso

NLP and LLM data scientist

Stuttgart, Germany
gloriagalasso • gloria-g-923657221
+39 351 838 5749 • fangalasso@gmail.com



ABOUT ME

NLP and LLM data scientist with an M.Sc. in Computational Linguistics and hands-on Python experience in text classification, information extraction and LLM fine-tuning. Experienced in building tested, reproducible ML pipelines on AWS, and in turning ambitious research ideas into working tools together with technical and domain experts.

WORK EXPERIENCE

06.2026–11.2026 **Clinical AI/ML Intern — Modality.ai**
San Francisco, USA (remote)
Multimodal LLMs for ALS assessment from speech and facial video
- Benchmarked 8 ML approaches on 7,389 patient videos; Qwen2.5-Omni embeddings outperformed clinical speech measures (balanced accuracy 0.72 vs. 0.66).
- Fine-tuned Qwen2.5-Omni-3B and explored LoRA and zero-shot prompting for video captioning.
- Built a tested Python pipeline (PyTorch, scikit-learn, pytest) on AWS with Git.

03.2026–09.2026 **Research Assistant — University of Stuttgart**
Computational Social Science Department — Stuttgart, Germany
- Built and evaluated named-entity extraction pipelines with GLiNER on biomedical patent claims and research papers; ran error analysis on missed entities and boundary errors.

2022–2023 **Project Coordinator — GOPA Com., Brussels, Belgium**
EU/public affairs & communications projects.

2020–2021 **Project Manager — WONCA Europe / EPR, Brussels, Belgium**
European healthcare communications & policy support.

NLP & SPEECH PROJECTS

2025–2026 **Text Classification for Alzheimer's Disease Detection — M.Sc. Thesis**
University of Stuttgart — Grade: 1.3
- Fine-tuned BERT-based transformers to classify spontaneous-speech transcripts, and compared them with logistic regression on linguistic features engineered with the ELFEN library.
- **Interpretability:** applied Integrated Gradients to identify which words drove the transformers' predictions, and compared them with the most predictive engineered features.
- Evaluated cross-dataset generalization and subgroup robustness.

2025 **Infant Cry Audio Classification**
- Extracted MFCC and log-Mel spectrogram features with Librosa and trained 1D CNNs (TensorFlow/Keras) for infant-cry classification.

TECHNICAL SKILLS

NLP & LLMs:	LLMs • Generative AI • Transformers (BERT, Qwen) • Text classification • Information extraction / NER • LoRA fine-tuning • Prompt engineering
ML:	Python • PyTorch • scikit-learn • Hugging Face Transformers
Speech/Audio:	Librosa • Praat • MFCC • Mel-spectrograms
Tools:	AWS • Git/GitHub • pytest • Poetry • Linux • LaTeX
Languages:	Italian (native) • English (C1) • German (B2, B2.2 in progress) • French (B1) • Spanish (B1)

EDUCATION

10.2023–07.2026 **Master of Science in Computational Linguistics**
University of Stuttgart — Stuttgart, Germany
- **Final grade:** 1.7 (German scale, 1.0 = best, 4.0 = lowest pass); M.Sc. thesis 1.3; research module 1.0.
- **NLP & ML coursework:** Topics in Natural Language Processing, Text Technology, Information Retrieval and Text Mining, Computational Semantics, Deep Learning for Speech and Text Processing.
- **Speech & health coursework:** Speech Analysis for Health Care Applications (1.3), Current Topics in Speech Technology (1.3), Topics in Phonetics (1.3).

09.2019–04.2022 **Master of Arts in International Development**
University of Kent, Brussels School of International Studies — Brussels, Belgium
- Graduated with Merit; coursework in international human rights law and health care in humanitarian disasters.

2016–07.2019 **B.A. in Linguistic Sciences for International Relations**
Università Cattolica del Sacro Cuore — Milan, Italy
- Final grade: 110/110 cum laude; English, German and Spanish, general linguistics.