

MARÍA GARCÍA-ABADILLO VELASCO

MSC. LANGUAGE ANALYSIS AND PROCESSING

Tools and languages

pandas • numpy • nltk • spacy • flair • matplotlib & seaborn • scikit-learn • pytorch & tensorflow • huggingface

Programming languages: Python • basic R • basic MATLAB

Natural languages: Spanish (Spain, native) • German C1 (TestDaF 20) • English C1 (CAE B)

Experience

Independent German language annotator

Toloka, Remote

February 2024 – Today

- Strictly followed annotation guidelines for German texts and & corrected grammatical errors

Student Research Assistant for Automatic Hate Speech Detection

Universität Bielefeld, Bielefeld

May 2023 – July 2023

- Generation of experimental stimuli for unimodal (text-based) and multimodal settings
- Conducted an eye-tracking experiment (SR Eyelink 1000 Plus) with more than 40 participants
- Corpus analysis for the identification of hate speech

Machine Translation Consultant

dSPACE, Paderborn

June 2022 – September 2022

- Developed a roadmap for researching, configuring, and evaluating five machine translation engines, culminating in a presentation and report detailing the final solution.

Open-source projects

- **LaSolanaNews Dataset:** Creation of a dataset with 1093 articles written over 14 years about the city of La Solana, design of annotation guidelines and manual annotation of instances. More information at [Kaggle](#)
- **LaSolana Embeddings:** Data collection on La Solana (≈ 4 M words) for the creation of static embeddings and their visualization in the TensorFlow Embedding Projector. More information is available at [GitHub](#)

Education

MSc. Language Analysis and Processing

University of the Basque Country, Donostia

September 2023 - Today

90 ECTS: Linear Algebra & Statistics, heuristic, statistical & neural methods for Computational Semantics & Morphology, Machine & Deep Learning, Synthesis & Speech Recognition, fine-tuning of LLMs for other applications: Lemmatize, NER, Sentiment Analysis, Argument Mining, Machine Translation or RAG

Short-study. Linguistics (Profile Computational)

Universität Bielefeld, Bielefeld

September 2022 – July 2023

30 ECTS: Evaluation of Conversational Agents, Intro. to Computational Linguistics, Hate Speech Detection, Computational Semantics, Neural Networks for NLP

BA. Translation and Interpreting

Universidad Complutense de Madrid, Madrid

September 2018 – June 2022

240 ECTS (36 with honors): Conference Interpreting, Consecutive & Bilateral, CAT tools, Specialized and General Translation DE/EN-ES. Erasmus scholarship at *Sprachen & Dolmetscher Institut* (2020-2021, Munich)

Publications

- *DeepL a prueba. Evaluación y análisis de la traducción al español del artículo Die eiserne Kanzlerin según la tipología de errores DQF-MQM* (2022). Bachelor Thesis with Honors (Matrícula de Honor)

Certificates

- **DeepLearning.AI:** Preprocessing Unstructured Data for LLM Applications (2024)
- **Codecademy:** Seven courses ranging from the command line to building machine learning models (2022/23)

Distinctions

- Banco Santander Erasmus scholarship for Erasmus students with the best grades (2021)
- Grant awarded by the Universidad Complutense de Madrid to students with the best academic results in the university entrance exams. (2018)