

Alberto Muñoz Ortiz

Lausanne, Switzerland | EU Citizen | alberto.munoz.ortiz@proton.me | github.com/amunozo

NLP Research Engineer

SUMMARY

Research engineer with a PhD in Computer Science, a background in Physics, and over five years of experience building language-model training and evaluation workflows. First author of 8 papers at ACL, EMNLP, and COLING.

Available immediately.

WORK EXPERIENCE

PhD Researcher University of A Coruña	Mar 2021 — Jun 2026 <i>A Coruña, Spain</i>
- Published 8 first-author papers on LLM evaluation, data-efficient NLP, structured prediction, and multilingual transfer. - Developed sequence-labeling methods for parsing and nested named entity recognition. - Built pipelines comparing human and LLM-generated text across syntactic, lexical, and emotional features. - Ran reproducible training and evaluation workflows on GPU clusters with Docker, Kubernetes, Slurm, Run:ai, and Weights & Biases.	
Visiting Researcher EPFL	Oct 2024 — Mar 2025 <i>Lausanne, Switzerland</i>
- Studied how transformers represent and generalize knowledge in Antoine Bosselut's NLP Lab. - Built Docker and Run:ai workflows to queue, monitor, and log parallel GPU runs.	
Visiting Researcher Ludwig-Maximilian University	Jul 2023 — Oct 2023 <i>Munich, Germany</i>
Built and trained a pixel-based language model in PyTorch for non-standardized German dialects. The work was published at COLING 2025.	
Python Developer Intern Alstom	Nov 2019 — Nov 2020 <i>Madrid, Spain</i>
Replaced a manual spreadsheet process with a Python and SQLite application. Added validation and reporting used by engineering stakeholders.	

EDUCATION

University of A Coruña <i>PhD in Computer Science</i>	A Coruña, Spain Oct 2021 — Jun 2026
<i>Thesis:</i> "Encoding, Representation, and Interpretation of Linguistic Structure Across NLP Paradigms".	
UNED <i>MSc in Language Technologies</i>	Spain 2020 — 2021
Complutense University of Madrid <i>BSc in Physics</i>	Madrid, Spain 2014 — 2019

SKILLS

- Research:** LLM evaluation, model behaviour, data-efficient NLP, structured prediction, information extraction, multilingual transfer, probing, error analysis
- ML stack:** Python, PyTorch, Hugging Face (Transformers, Datasets), scikit-learn, pandas, NumPy, SciPy, spaCy, NLTK
- Infrastructure:** Docker, Kubernetes, Slurm, Run:ai, Git, Linux, pytest, Weights & Biases, SQLite
- Languages:** Spanish (Native), English (C2), French (B2), German (B1), Chinese (A1)