

Dr. Abelardo Carlos **Martinez Lorenzo**

PH.D. IN DEEP LEARNING AND NATURAL LANGUAGE PROCESSING

☎ (+34) 653903060 | ✉ abelardomartinezlorenzo@gmail.com | 🏠 carlosml26.github.io | 📺 CarlosML26 | 📺 carlosml26 | 🎓 Abelardo Carlos

Experience

Lead AI Scientist (Team Lead + Senior Research Engineer)

Washington, DC, USA

THE WORLD BANK (ETC)

Sept. 2024 – Current

- Led and supervised the AI science team of ImpactAI, with 6 AI Scientist. Objective: develop a cutting-edge **agentic framework** that **synthesizes evidence from RCTs** to deliver high-quality **meta-analyses** for policymakers.
- Spearhead the design and coordination of pipelines for **PDF parsing, information extraction, entity linking, SQL parsing, and summarization**, enabling scalable analysis of development economics literature.
- Manage the Google Cloud Platform (GCP) infrastructure, ensuring secure, scalable, and low-latency deployment of AI services.
- Fine-tuned LLMs (Gemini, GPT, LLaMA 3 (8B)) to create a domain-specific LLM for development economics, incorporating custom alignment and pretraining strategies. Developed a semi-automated taxonomy generation and refinement pipeline, combining LLM capabilities with human-in-the-loop validation.
- Represented ImpactAI at Google.org Impact Summit in Palo Alto, showcasing successful deployments and impact.
- Secured long-term funding by authoring successful proposals for the World Bank's KCP and other internal initiatives.
- Serve as an AI advisor for multiple World Bank projects, providing strategic guidance on LLM integration, data pipelines, and responsible deployment practices.
- One first-author paper under review at EMNLP 2025.

Senior AI Scientist

Washington, DC, USA

THE WORLD BANK (STC)

Jan 2024 – Sept. 2024

- Designed and fine-tuned a **domain-specific entity linking model** (bi-encoder using DeBERTa and e5) for mapping text to an impact evaluation taxonomy.
- Developed and maintained production-grade NLP services using Docker and FastAPI, deployed on **Google Cloud Platform (GCP)**.
- Engineered and evaluated **LLM prompting pipelines** (GPT-4o, LLaMA 7B, Gemini 1.5 Pro) for information extraction and semantic parsing.
- Led GCP resource management, including container orchestration, API deployment, and latency optimization.

NLP Researcher

Rome, Italy

SAPIENZA NLP - SAPIENZA UNIVERSITY

2023 – 2024

- Led NLP research projects with PhD students, focusing on semantic parsing and multilingual NLU.
- Developed and fine-tuned **LLM-based pipelines** (e.g., GPT-4o) for semantic parsing and structured annotation.
- Created efficient **encoder-decoder models** and coordinated the construction of a novel multilingual semantic parsing dataset.
- Built production-ready pipelines for **NLU tasks** including parsing, alignment, and tagging.
- Served as Teaching Assistant for MSc courses in NLP; mentored students in research methods and implementation.
- Published and presented three papers at top-tier conferences: **ACL, NAACL, and COLING**.

NLP Researcher

Rome, Italy

BABELSCAPE

2020 – 2023

- Designed a cross-lingual semantic representation framework applicable to multiple languages.
- Developed state-of-the-art **AMR parsing and ensemble models** using BART and T5.
- Built an **alignment-aware encoder-decoder training pipeline** using cross-attention signals in the loss function.
- Delivered Dockerized demos for multilingual semantic parsing and model inference.
- Published and presented five papers at **EMNLP, ACL, and AACL**.

Research Assistant

Berlin, Germany

TU-BERLIN

2018 – 2020

- Contributed to the design of a hybrid framework that combines knowledge graph embeddings with ontology-based reasoning for improved link prediction.
- Assisted in implementing a loosely coupled architecture allowing plug-and-play use of multiple KGE models (e.g., TransE, ComplEx) and OWL2-compatible reasoners.
- Supported experiments showing up to 3.5× improvement in prediction accuracy (MRR) over baseline KGE methods.
- Co-authored the publication "Towards Loosely-Coupling Knowledge Graph Embeddings and Ontology-based Reasoning"

Software Engineer

Malaga, Spain

FREELANCER

2017 – 2018

- Built a real-time multiplatform app for emergency team coordination with chat, voice, and live maps.
- Developed backend with Node.js (LoopBack) and MongoDB, and frontend with AngularJS.
- Integrated WebRTC for peer-to-peer voice communication to reduce latency.
- Improved response times by enabling fast incident creation and team assignment.
- Helping to organize ISCOG 2017 at Malaga's university

Education

Ph.D. in Deep Learning & Natural Language Processing (Marie Skłodowska-Curie)

Rome, Italy

SAPIENZA UNIVERSITY OF ROME

- Grant: Marie Skłodowska-Curie grant
- Project: EU-Horizon 2020 - KnowGraphs
- Thesis: Enhancing Semantic Parsing in the Age of Pre-trained Language Models
- Advisor: Roberto Navigli

MSc in Big Data Management and Analytics (BDMA)

Berlin, Barcelona, Brussels

MASTER

- Final Grade: **9.0/10**
- Relevant Coursework:
 - Technische Universität Berlin (**Tu-Berlin**): Large Scale Analytics
 - Universitat Politècnica de Catalunya (**UPC**): Big Data Fundamentals
 - Université Libre de Bruxelles (**ULB**): Business Intelligence Fundamentals

MSc in Big Data Management and Analytics (BDMA)

Berlin, Barcelona, Brussels

ERASMUS MUNDUS JOINT MASTER DEGREE PROGRAMME

- Final Grade: **9.0/10**
- Relevant Coursework:
 - Technische Universität Berlin (**Tu-Berlin**): Large Scale Analytics
 - Universitat Politècnica de Catalunya (**UPC**): Big Data Fundamentals
 - Université Libre de Bruxelles (**ULB**): Business Intelligence Fundamentals

BSc in Software Engineering

Málaga, Spain

UNIVERSIDAD MÁLAGA

- Final Grade: **9.0/10**
- Erasmus Exchange: Univerzita Karlova in Prague, Czech Republic

Relevant Publications

| | | |
|------|--|-------------|
| 2024 | Abelardo Carlos Martínez Lorenzo , Pere-Lluís Huguet Cabot, Roberto Navigli. Mitigating Data Scarcity in Semantic Parsing across Languages with the Multilingual Semantic Layer and its Dataset. | ACL 2024 |
| 2024 | Simone Conia, Edoardo Barba, Abelardo Carlos Martínez Lorenzo , Pere-Lluís Huguet Cabot, Ricardo Orlando, Luigi Procopio, Roberto Navigli. MOSAICo: a Multilingual Open-text Semantically Annotated Interlinked Corpus. | NAACL 2024 |
| 2024 | Abelardo Carlos Martínez Lorenzo , Roberto Navigli. Efficient AMR Parsing with CLAP: Compact Linearization with an Adaptable Parser. | COLING 2024 |
| 2023 | Abelardo Carlos Martínez Lorenzo *, Pere-Lluís Huguet Cabot*, Roberto Navigli. Cross-lingual AMR Alignment: Paying Attention to Cross-Attention. | ACL 2023 |
| 2023 | Abelardo Carlos Martínez Lorenzo *, Pere-Lluís Huguet Cabot*, Roberto Navigli. AMRs Assemble! Learning to Ensemble with Autoregressive Models for AMR Parsing. | ACL 2023 |
| 2023 | Pavlo Vasylenko*, Abelardo Carlos Martínez Lorenzo *, Pere-Lluís Huguet Cabot*, Roberto Navigli. Incorporating Graph Information in Transformer-based AMR Parsing. | ACL 2023 |
| 2022 | Abelardo Carlos Martínez Lorenzo , Marcu Maru, Roberto Navigli. Fully-Semantic Parsing and Generation: The BabelNet Meaning Representation. | ACL 2022 |
| 2022 | Roberto Navigli, Rexhina Blloshmi, Abelardo Carlos Martínez Lorenzo . BabelNet Meaning Representation: A Fully Semantic Formalism to Overcome Language Barriers. | AAAI 2022 |

Skills

| | |
|-----------------------------|--|
| Frameworks | Deep Learning & ML: PyTorch, PyTorch Lightning, TensorFlow, scikit-learn, Transformers, HuggingFace, WandB |
| LLM Ecosystem | LangChain, OpenAI API, Gemini API, LLaMA, Mistral, vLLM, RAG, instruction tuning, prompt engineering |
| Back-End Development | FastAPI, Django, RESTful APIs, Streamlit, Argilla, Uvicorn, Gunicorn |
| Cloud & DevOps | Google Cloud Platform (GCP), Docker, Cloud Run, Cloud Build, Cloud Functions, Pub/Sub, Vertex AI, BigQuery |
| Languages | Python, Java, SQL, Bash/Shell |
| Developer Tools | Git, GitHub, VS Code, Jupyter, Make, Poetry, PyPI, Tmux, Linux |
| Spoken Languages | English (fluent), Spanish (native), Italian (fluent), Portuguese (intermediate) |

Honors & Awards

- 2023 **Winner**, 2× Italian Supercomputing Resource Allocation (ISCRA) Call for Proposal Type C – compute hours at CINECA HPC Center *Rome, Italy*
- 2023 **Award**, Best Reviewer Award – ACL *Toronto, Canada*
- 2020 **Grant**, Marie Skłodowska-Curie PhD Fellowship – Horizon 2020 Project *Rome, Italy*

Program Committees

2020–2025 **Program Reviewer**, ACL Rolling Review (ARR): reviewer

2024 **Reviewer**, ARR February 2024, ARR June 2024

2023 **Reviewer**, ARR December 2023, ARR October 2023, EMNLP 2023, ACL 2023, EACL 2023

2022 **Reviewer**, LREC 2022

2021 **Reviewer**, EMNLP 2021