UMBERTO SELVA

NLP / AI Engineer, Linguistics PhD

@ u.o.selva@gmail.com

🔗 linkedin.com/in/umberto-selva

🔗 github.com/umbertoselva

📞 +39 3335218014

📍 Italy / Netherlands

✎ English, Italian

PROFILE

I am a former post-doctoral researcher and university lecturer with a PhD in Linguistics from Leiden University in the Netherlands. After a successful career in academia, I trained and freelanced as an NLP specialist, acquiring a broad expertise across Machine Learning and Data Science, with a solid foundation in Software Development. I currently work as an Artificial Intelligence Engineer, and I specialize in designing LLM-powered solutions and Prompt Engineering for a web platform in the pharmaceutical sector. I continue to provide freelancing services.

EXPERIENCE

Updated: Jan. 2025

• AI Engineer

Pharmazie.com   📅 **2023–today**

Pharmazie.com is a Germany-based web platform that gathers data from over 25 databases to provide reliable info on pharmaceutical products to healthcare professionals. I oversaw the company's transition to Generative AI (GenAI), single-handedly implementing projects end-to-end, such as LLM-powered conversational assistants, APIs, web apps, and pipelines, designing all step from text/table data extraction, vectorization/embedding, retrieval augmented generation (RAG), AI agents and tools, fine-tuning, evaluation, prompt engineering, and (intent/topic) conversation design, all the way to server management and production deployment.

Skills: Python, NLP, LLMs, OpenAI, GPT, LangChain, LangGraph, LlamaIndex, Conversation Design (Voiceflow), Prompt Engineering, Chatbots, Retrieval Augmented Generation (RAG), Ragas, AI Agents, Fine-Tuning, API, Back-end Web Dev, Unstructured, AzureOpenAI, VertexAI (Google Cloud), Kubeflow, Azure and AWS Cloud.

• NLP / AI / ML / Data Science Consultant

Freelance   📅 **2020–today**

Between 2020-2022 I transitioned from academia to the private sector, training in NLP, AI, Machine Learning/ Deep Learning and Data Science, and I have since been providing freelancing services.

Skills: Python, NLP (NLTK, spaCy, Haystack,   etc.), Conversation Design (Voiceflow, Botpress), ML/DL (Tensorflow, Keras, Scikit-Learn, Huggingface), Data Science (Pandas, NumPy, Matplotlib, etc.), Statistics, Web Development (FastAPI, Flask, Streamlit, etc.), SQL (MySQL, SQLite), bash / zsh, Linux, Jupyter, git, Docker

• Academic Researcher and Lecturer (Linguistics)

LUCL / IIAS   📅 **2014–2020**

Between 2014-2020, I was research assistant, then PhD researcher, post-doctoral researcher and lecturer at the Leiden University Centre for Linguistics (2014-19) and the International Institute for Asian Studies (2019-20).

Skills: Linguistics (General, Historical, Computational), Academic Research, Academic Writing, Academic Publishing, Teaching, Research Project Management

EDUCATION AND PRIZES

PhD Linguistics (Leiden Uni, 2019) • Research MA Linguistics (Leiden Uni, 2014, grade: 9) • BA Classics (University of Turin, 2011, grade: 110/110 *cum laude*) • 2015 First Prize for Best MA Thesis by the Indogermanische Gesellschaft.

CERTIFICATIONS

Machine Learning (Stanford, 2020) • NLP Specialization (DeepLearning.ai = DL.ai, 2022) • NLP with Transformers in Python (Udemy, 2022) • Deep Learning Specialization (DL.ai, 2022) • Statistics for Data Science with Python (IBM, 2022) • Practical Data Science on the AWS Cloud (AWS, DL.ai, 2022) • Fine Tuning LLMs (Lamini, DL.ai, 2024) • Building Agentic RAG (LlamaIndex, DL.ai, 2024) • LLMOps (Google Cloud, DL.ai, 2024) • Functions, Tools and Agents with LangChain (LangChain, DL.ai, 2024)