Felipe Espinosa

Paris, France | +33 7 66 09 03 35 | felipeespinosa.13@outlook.com | `https://felipe-espinosa.vercel.app`
`https://linkedin.com/in/felipe-espinosa13` | `https://github.com/Pipe1213` |

SUMMARY

AI/ML Engineer with expertise in agentic, retrieval-augmented, and LLM-powered systems. Developed a functional SQL Agent prototype at CFM using **LangChain** and **LangGraph** for natural-language data querying. Also developed the first Walloon **TTS** system published at SSW 2025, applying Gen-AI to solve real data challenges.

EXPERIENCE

Capital Fund Management (CFM) | Paris, France

AI Engineer Intern   March 2025 – September 2025

- Achieved **88% accuracy** and **93% correct table retrieval** on an internal test set built with the data team.
- Designed a **hybrid search engine** using **BM25**, **semantic embeddings**, and **Reciprocal Rank Fusion (RRF)** to improve table relevance and precision across **Oracle** databases.
- Created **RAG pipelines** with schema enrichment, multi-step reasoning, and fallback strategies to make the system more reliable for complex financial queries.
- Built the complete solution including retrieval modules, validation tools, an evaluation pipeline, and a **Streamlit** interface tested and adopted by internal analysts.
- Delivered a **functional prototype** that now serves as the foundation for future **LLM-based** automation tools.

Laboratoire Interdisciplinaire des Sciences du Numérique (LISN) | Orsay, France

NLP Research Intern   February 2024 – August 2024

- Trained multiple **VITS** models with and without French pretraining, achieving a **Mean Opinion Score (MOS)** of **4.22** for the best configuration.
- Conducted objective evaluations using **PESQ**, **MCD**, and **SECS** metrics, and validated the results through perceptual tests with native speakers.
- Published results at the **13th Speech Synthesis Workshop (SSW 2025)**, highlighting cross-lingual transfer strategies for low-resource **TTS**.
- Released trained models, demos, and inference scripts publicly on **Hugging Face** to support open research and community use.

SKILLS

Programming & Databases: Python, SQL, Oracle, Redis, ChromaDB
AI & ML Frameworks: PyTorch, LangChain, LangGraph, Transformers, Hugging Face, Scikit-learn, Pandas, OpenCV, NumPy, Matplotlib, Deep Learning, Generative AI, Agent Architectures, RAG, Hybrid Search, Fine-tuning, Prompt Engineering, Vector Databases, Model Deployment
Tools & Platforms: Git, Docker, Slurm, AWS, Streamlit, Jupyter, Confluence, Jira
Soft Skills: Analytical Thinking, Problem Solving, Teamwork, Communication, Adaptability, Autonomy
Languages: Spanish (Native), English (C1), Portuguese (B2), French (B1)

PROJECTS

Creation of a Synthetic Dataset for French NER – Project link
Developed a multilingual pipeline using Hugging Face Transformers and **fine-tuned** XLM-RoBERTa for cross-lingual NER in clinical trial texts.

Structure from Motion 3D Reconstruction – Project link
Implemented a full SfM pipeline with SIFT, FLANN, and PnP to reconstruct 3D scenes from multiple 2D images using only RGB cameras.

EDUCATION

Paris Saclay University   September 2023 – September 2025
Master Degree in Artificial Intelligence

Lloyds International College   February 2020 – February 2022
Advanced Diploma in Leadership and Management

Universidad Industrial de Santander   April 2010 – December 2016
B.Sc. in Electronic Engineering