



Duarte Fernández Piñeiro

Artificial Intelligence Engineer | Graduate, University of Vigo (2026)

Santiago de Compostela, A Coruña |
[Interactive Portfolio](#) | English B2

| dfernandezpineiro@gmail.com | [LinkedIn](#) | [GitHub](#) |

PROFILE

Artificial Intelligence graduate with hands-on experience in generative AI, RAG, automation, and analytics. I develop end-to-end Python solutions: document processing, evidence retrieval, APIs, automated testing, and cloud deployment.

EXPERIENCE

AHORA Enterprise | Artificial Intelligence and Data Analytics Internship | Mar. 2026 - Jun. 2026

- Contributed to the development of a RAG solution for querying corporate documentation through natural language, from document processing to the query interface.
- Contributed to evidence retrieval and grounded answers; validated the solution with AHORA staff using real documentation.
- Developed Power BI dashboards, supported the migration from Qlik, and researched semantic models and local alternatives for LLM-based solutions.

FEATURED PROJECTS

Professional Interview Assistant | [Demo](#) | Python, FastAPI, OpenAI Responses API, JavaScript, Docker, DigitalOcean

- Public web application for querying my profile through natural language, with dynamic document context and streaming responses. Implemented sessions, rate limiting, security, concurrency control, and Docker deployment; validated with 18 automated tests.

Hybrid RAG System for Corporate Documentation | [Explainer Video](#) | Bachelor's Thesis | Python, BM25, ChromaDB, RRF, Jina Reranker, PyMuPDF, Pydantic

- Designed a modular architecture for querying PDFs with grounded answers and traceability down to document, section, and page. Combined lexical and semantic search with reranking; validated it using 109 documents, 2,913 traceable chunks, and 57 automated tests.

TwinPhoto / Similar Photos V3 | [Download](#) | Python, NumPy, SQLite, Tkinter, Rust, Windows Portable Devices

- Desktop application for detecting and grouping visually similar images on computers and USB-connected mobile devices. Implemented 64-bit perceptual hashing, parallel processing, incremental SQLite indexing, and a native Rust bridge for verified batch synchronization and deletion.

Bitcoin Sentiment Classifier | [Demo](#) | Python, Sentence Transformers, LightGBM, scikit-learn, FastAPI

- Trained and evaluated a binary classifier for posts using embeddings and LightGBM. After optimizing the decision threshold, it achieved 87.5% accuracy and 88.1% F1 on the test set; integrated it into an explanatory API.

TECHNICAL SKILLS

Generative AI and retrieval: LLMs, RAG, OpenAI API, embeddings, BM25, ChromaDB, Jina Reranker, LangChain, Ollama.

Machine learning and data: scikit-learn, PyTorch, LightGBM, Sentence Transformers, pandas, NumPy, OpenCV, Power BI.

Engineering and cloud: Python, SQL, Java, JavaScript, FastAPI, REST APIs, Git, Docker, Linux, DigitalOcean, automated testing.

Duarte Fernández Piñeiro

Detailed CV

Artificial Intelligence Engineer

GenAI, RAG, NLP, LLMs, Agents, Codex and Software Development

Residence: Santiago de Compostela, A Coruña

Email: dfernandezpineiro@gmail.com

Phone:

LinkedIn: linkedin.com/in/dfernandezpineiro

GitHub: github.com/DuarteFernandezPineiro

Personal Assistant App for Interviewers: [Link](#)

Professional Profile

Artificial Intelligence graduate with internship experience in applied AI, automation, and data analytics. I specialize in turning information needs into complete technical solutions using Python, generative AI, NLP, and RAG: from data processing and evidence retrieval to API integration, automated testing, and cloud deployment.

I combine machine learning foundations with a practical software engineering mindset. I have worked with real corporate documentation, validated solutions with users, and built a public conversational application so that interviewers can explore my profile in greater depth. I bring technical curiosity, autonomy to research and build prototypes, and attention to traceability, maintainability, and user experience.

Professional Experience

AHORA Enterprise | Artificial Intelligence, Data Analytics and Sales Internship

March 2026 - June 2026

- Worked in a multidisciplinary team on the development of a RAG solution for querying corporate documentation through natural language.
- Contributed to document processing, evidence retrieval, grounded answer generation, and the query interface; validated the solution with AHORA staff using real documentation.
- Participated in AI automation initiatives for topic assignment and the analysis, detection, and classification of invoice information. This experience allowed me to apply information-processing techniques to specific operational needs.
- Developed Power BI dashboards and participated in migrating visualization projects from Qlik. I also researched reusable semantic models and local, open-source alternatives to reduce cost and improve the feasibility of LLM-based solutions.

Selected Projects

Professional Interview Assistant | Personal cloud-deployed project

[Link](#)

Python, FastAPI, OpenAI Responses API, JavaScript, Docker, DigitalOcean

Public web application designed to allow recruiters and interviewers to explore my education, experience, and projects through natural-language questions. The assistant answers from controlled documentation, turning a static CV into a functional demonstration of how I approach a generative AI product.

- Designed and developed the end-to-end solution: FastAPI/ASGI backend, JavaScript interface, dynamic document context selection, and streaming responses using OpenAI Responses API.
- Implemented server-side sessions, rate limiting, security headers, and concurrency control with 2 simultaneous generations and up to 20 requests in a FIFO queue. These decisions aim to maintain a reliable experience and control use of the public application.
- Containerized the service with Docker, deployed it on DigitalOcean, and added optional consent-based analytics. Validated the main behaviors with 18 automated tests.

Hybrid RAG System for Corporate Documentation | Bachelor's Thesis

[Explainer Video Link](#)

Python, OpenAI Responses API, OpenAI embeddings, PyMuPDF, Pydantic, BM25, ChromaDB, RRF, Jina Reranker

Academic prototype for querying corporate PDF documents through natural language, with grounded answers and traceability to sources. The project demonstrates my ability to design a RAG architecture beyond a simple language-model call.

- Designed and implemented a modular architecture with an offline pipeline for page-level extraction, cleaning, metadata enrichment, chunking, and indexing, and an online pipeline for retrieval, reranking, and source-backed answer generation.
- Combined BM25 lexical retrieval and semantic search in ChromaDB using Reciprocal Rank Fusion. Added Jina Reranker to prioritize relevant evidence and maintained traceability to the original document, section, and pages.
- Handled PDF heterogeneity with validation using Pydantic and persistent artifacts to reconstruct stages and selectively reindex. Added a CLI, local web interface, and 57 automated tests across 109 documents and 2,913 traceable chunks.

Bitcoin Sentiment Classifier

[Link](#)

Python, Sentence Transformers, LightGBM, scikit-learn, FastAPI

Application that analyzes Bitcoin-related posts and turns sentiment predictions into an understandable explanation. This project strengthens my ability to build and evaluate machine-learning pipelines that connect a model to a user experience.

- Designed, trained, and evaluated a binary tweet classification pipeline using semantic embeddings and a LightGBM ensemble, applying stratified cross-validation and metric analysis.
- Optimized the decision threshold on the validation set and achieved approximately 87.5% accuracy and 88.1% F1 on the test set. Integrated the model into an API to explain the detected sentiment cautiously.

TwinPhoto / Similar Photos V3

[Download Link](#)

Python, NumPy, SQLite, Tkinter, Rust/WPD

Designed and developed a desktop application to safely detect, group, and delete visually similar images in local folders and galleries on mobile devices connected via USB.

- Implemented 64-bit perceptual hashing, parallel processing, and persistent SQLite indexing to avoid quadratic comparisons and recalculate only modified files, enabling efficient analysis of large collections.
- Integrated a native Rust bridge over Windows Portable Devices for incremental synchronization and verified batch deletion. Completed the product with an interactive interface, automated tests, and installable packaging for Windows.

Knowledge and Skills

My profile combines technical depth with execution capability: I can reason about information retrieval, embeddings, and model evaluation, and also turn those components into APIs, web applications, and reproducible deployments. I am used to documenting technical decisions, testing critical behaviors, and explaining complex solutions clearly. This combination allows me to contribute both to the exploration of an AI solution and to its practical construction and validation.

I also have experience working on and developing all kinds of software with support from LLMs such as GPT or Claude, together with prompt engineering, to achieve better and faster results.

Additional academic foundation: symbolic AI and algorithms (MiniMax, alpha-beta pruning, and BDI agents), NLP from first principles (BPE and Skip-gram), computer vision and deep learning, and semantic data/knowledge graphs (RDF, RDFS, SPARQL, Apache Jena, and Fuseki).

Education

[Bachelor's Degree in Artificial Intelligence | University of Vigo, ESEI](#)
[September 2022 - June 2026](#)

Coursework in machine learning, deep learning, NLP, information retrieval, computer vision, data analysis, databases, software engineering, distributed systems, and symbolic AI. The program is complemented by practical projects focused on building, evaluating, and explaining AI solutions.

Technical Skills

- AI engineering: end-to-end pipeline design, document processing, modular architecture, information retrieval, API integration, automated testing, technical documentation, and experimental evaluation.
- Generative AI and NLP: LLM, RAG, OpenAI API, embeddings, BM25, RRF, ChromaDB, Qdrant, Jina Reranker, LangChain, LlamaIndex, Ollama, Sentence Transformers, and Transformers.
- Machine learning and data: scikit-learn, PyTorch, TensorFlow, LightGBM, pandas, NumPy, OpenCV, Power BI, Qlik, and Jupyter.
- Development and cloud: Python, SQL, Java, R, FastAPI, Flask, REST APIs, JavaScript, Git, Docker, Linux, DigitalOcean, and Google Cloud.
- Generative AI models: Claude, Codex, Agents, skills, prompt engineering.

Interests and Strengths

I am a sociable, positive, and solution-oriented person, with an ability to connect with others, adapt to different profiles, and work as part of a team. I approach problems calmly, staying focused and looking for practical solutions.

Sport is an important part of my life. I compete in athletics and practice cycling, surfing, hiking, and other outdoor activities. This consistency in sport has strengthened qualities that I also bring to the professional environment, such as discipline, perseverance, concentration, and decision-making in demanding situations.

Languages

Native Spanish and Galician | English B2 | French B1 (DELF certification)