

Dinorah Farias

Portfolio | dinorahfarias@outlook.com | linkedin.com/in/dinorahfariasc | github.com/dinorahfariasc | Lattes

SUMMARY

AI and back-end developer in Python with a soft spot for data science. I build systems that run in production: multi-agent architectures, internal tools, and document extraction pipelines with LLMs and OCR. My background mixes applied AI research with industry work, and I care about measuring results, validating model output, and keeping costs under control. I use spec-driven development and coding agents in my daily workflow.

EXPERIENCE

Back-end Developer Junior III - Lightfarm Studios

Feb 2026 - Present

Python, Multi-agent Systems, OpenRouter, Self-hosted LLMs

- Built an agent that recommends work assignments by seniority, requirement fit, vacations, and similarity to past work.
- Designed the multi-agent architecture behind an internal chatbot and automated operational-capacity dashboards.
- Deployed self-hosted models on VMs in production; development with Spec Kit (SDD) and Orca.

Junior Python AI Developer - LANCE/IMD UFRN

Nov 2025 - Present

Python, FastAPI, LLMs, RabbitMQ, Tesseract

- On the YVY Energia project, built an LLM hub that extracts utility-bill data from any provider and layout, with no templates or manual annotation.
- Rule-based validation (check digits, totals, taxes) with auto-correction, provider fallback, and versioned prompts.
- On real bills: consumption accuracy 75% → 100%, max time per bill −82%, cost per bill −54% or more.

AI Research and Development Fellow - Public Prosecutor's Office of RN

Jun 2025 - Feb 2026

Python, FastAPI, PyTesseract, Azure, Elasticsearch

- Cut OCR extraction of 500+ page legal PDFs from 15 to 3 minutes; GPU-parallel OCR for a ~200% speedup.
- Fed extracted text into embeddings for an AI pipeline that pre-analyzes each legal case.
- Benchmarked Cassandra, MongoDB, and PostgreSQL (vector) for search workloads; evaluated DeepSeek on Ollama.

Scientific Initiation Researcher - CNPq, UFRN/NUT

May 2025 - Nov 2025

Python, Generative AI, FastAPI, MongoDB, PostgreSQL

- On the RHAЕ project, built an AI assistant and agents for production planning (PPCP) in small textile companies, using production and IoT data.

PROJECTS

Pintado (LangGraph, MCP, FastAPI, React): multimodal orchestrator agent served as an MCP server in microservices.

NOAH (Python, LangChain, Ollama): emergency chatbot for natural disaster victims, Scientific Initiation 2025.

TECHNICAL SKILLS

AI & LLMs: LangGraph, LangChain, Multi-agent Systems, RAG, MCP, OpenAI, Google AI, OpenRouter, Ollama, PyTorch

Documents: Tesseract, Azure, PyMuPDF, Docling, Label Studio

Back-end & Data: Python, FastAPI, RabbitMQ, Docker, PostgreSQL, MongoDB, Elasticsearch, Pandas, Go

Workflow: Spec Kit, spec-driven development (SDD), Orca, code review

EDUCATION

Federal University of Rio Grande do Norte - UFRN

Mar 2023 - Jul 2026

Interdisciplinary Bachelor of Information Technology

Federal Institute of Rio Grande do Norte - IFRN

2019 - 2022

Technical High School in Digital Game Programming

ADDITIONAL

Publication: Systematic Review of PPCP Models, Artificial Intelligence, and Optimization Applied to the Small-Scale Textile Industry (Springer)

Certifications: Designing Agentic Systems with LangChain; Data Analysis with Python (IBM); Python for Data Science and AI (IBM)

Languages: Portuguese (native), English (advanced), French (basic)