Henrique Pimenta Souza

Python | AI Automation | Machine Learning | RAG | Data Structures | PyTorch | LLM | LangChain

São Paulo, Brazil
Email: pimenttasouzza02@gmail.com | Phone: +55 11 98110-4828
LinkedIn: linkedin.com/in/henriquepimentaofc777777777

Professional Summary

Results-driven Backend Developer with a robust foundation in Python, FastAPI, PostgreSQL, and Generative AI. Specialized in building high-performance APIs, intelligent agents using LLMs, and RAG pipelines. Proficient in modern software architecture (DDD, microservices), data engineering workflows, and real-time observability. Passionate about AI automation, prompt engineering, and backend systems infused with LLM capabilities.

Technical Skills

Backend: Python, FastAPI, .NET Core (C#), Java (Spring Boot), RESTful APIs, Domain-Driven Design (DDD)
Databases: PostgreSQL, MySQL, MongoDB, SQLite
AI & Automation: OpenAI, Gemini, LangChain, LangSmith, LangGraph, RAG, Whisper, OCR (EasyOCR), Prompt Engineering
DevOps: Docker, Git, CI/CD, AWS
Observability: Grafana, Prometheus, Loki
Frontend (Basic): Angular, Vue.js
Tools: Postman, Insomnia, Swagger, VS Code
Methodologies: Scrum, Agile, BPMN, C4 Model, UML
Languages: Portuguese (Native), English (Professional Working Proficiency)

Professional Experience

Full-Stack Developer — *Stefanini Brasil*
São Paulo, Brazil | *Aug 2024 – Present*

- Developed scalable enterprise systems in Python (FastAPI) and .NET Core with modular, testable architecture

- Designed and maintained robust RESTful APIs integrating PostgreSQL and MongoDB

- Automated business intelligence data flows using Prometheus and Grafana

- Built intelligent agents with OpenAI and Gemini APIs via LangChain for summarization and QA

- Mapped business processes using BPMN and produced C4 diagrams for architectural documentation

---

Key Projects

RAG API for Curriculum Analysis (*Desafio Técnico - TheGarage*)
Developed an intelligent backend using FastAPI, PostgreSQL, and LangChain. The system parses curricula (PDF/image), performs OCR (EasyOCR, Whisper), and applies Gemini-based RAG pipelines for answering business queries. Modular and optimized for reliability and traceability via LangSmith.

OCR + LLM API for Intelligent Resume Analysis
Production-ready RESTful API built with FastAPI, PaddleOCR, and Hugging Face Transformers for automated resume extraction and interpretation. Supports PDF/image OCR, LLM-based summarization and Q&A, and logs requests in MongoDB for traceability. Fully containerized with Docker and deployable via Docker Compose. Emphasizes modular architecture, environment-based configuration, and structured logging.
Tech Stack: FastAPI · Python · PaddleOCR · Hugging Face Transformers · PyMongo · MongoDB · Docker · Docker Compose · RESTful API Design

Jira Automation Platform
Built a real-time backend pipeline that extracts Jira issue data and streams insights to Grafana. Used FastAPI with async tasks and PostgreSQL. Integrated OpenAI Whisper for audio transcription workflows.

Sai App
Architected the microservice backend of an AI-powered productivity app. Focused on clean DDD patterns, asynchronous operations, and secure API boundaries.

DevRequest
Engineered backend for AI-based candidate matching platform. Applied GPT for intelligent candidate scoring. Stack: Python, MongoDB, FastAPI.

Mia — AI Assistant
Developed backend infrastructure for context-aware empathetic assistant. Emphasis on vector-based memory and LLM personalization.

Education

Bachelor of Technology in Systems Analysis and Development
São Paulo Tech School (SPTech) | Jan 2024 – Jun 2026

Technical Degree in Database Systems
ETEC Ferraz de Vasconcelos | Jan 2020 – Dec 2021

Certifications

- *Foundations of Project Management* (Google Career Certificates)

- *Asking Questions to Make Data-Driven Decisions*

- *Data Cleaning Techniques*

- *Initiating a Successful Project*

- *Preparing Data for Exploration*