

PEDRO REICHOW

AI Platform Engineer • RAG & Agentic Systems • Technology Entrepreneur

Santa Catarina, Brazil • pedro_reichow@hotmail.com • linkedin.com/in/pedroreichow • [Talk to My Resume](#)

PROFESSIONAL SUMMARY

AI Platform Engineer combining production-grade LLM systems with robust software engineering. Expert in **Hybrid Retrieval** (Vector + Knowledge Graphs via **Cognee**), **RAG Evaluation Harnesses**, and multi-tool agent orchestration. 7+ years building scalable AI platforms with focus on **observability (SLOs, p50/p95 latency)**, security, and cost optimization. Proven track record delivering enterprise systems that transform manual processes from weeks to minutes. Fluent in English, intermediate Spanish.

PROFESSIONAL EXPERIENCE

Senior Software & Data Engineer | AI Observability CVS Health

Aug 2024 – Present

- Architected **SLOs for AI pipeline latency (p50/p95)** and cost-per-token monitoring using Datadog/Grafana for healthcare AI systems in **HIPAA/FedRAMP** compliance.
- Built observability dashboards tracking **faithfulness and relevance metrics** for LLM outputs, enabling prompt version A/B testing and continuous improvement.
- Full-stack development of enterprise AI applications with predictive analysis and workflow automation.

Cloud AI Solutions Architect Santodigital (Google Partner)

Jan 2024 – Jul 2024

- Architected **production-grade RAG with BigQuery + Vertex AI**, reducing document analysis from weeks to minutes with offline evaluation suites for retrieval quality.
- Implemented **reinforcement loops via ADK**: observability logs feeding prompt fine-tuning, reducing hallucinations and improving faithfulness scores in production.
- Built AI Agents for enterprise clients: compliance automation, contract alerts, auto-service with **Zendesk handoff** and human-in-the-loop for critical actions.

Senior Solutions Consultant (Solutions Consultant IV) Capgemini Brasil

Jan 2022 – Dec 2023

- Architecture of **Knowledge Graph** solutions and data engineering for fraud detection and risk analysis in financial sector.
- Development of sophisticated ontologies using **Protégé and RDF/OWL** for complex business intelligence and semantic search.
- Technical collaboration with teams of **50+ engineers** on enterprise data engineering optimizations.

Full Stack Engineer | VR & Real-Time Systems Simulated Reality

Jan 2021 – Dec 2021

- Developed VR solutions in **Node.js and React** for training and wellness applications with 3D object interaction and physics simulations.
- Built scalable APIs with queues for asynchronous processing and **pub/sub for real-time communication** between system components.
- Engineered pipelines for processing large 3D video files and game updates compiled in **Unity and Unreal Engine**.

Senior Data Engineer QI Tech (Fintech)

Jan 2019 – Dec 2020

- Processed **500GB+ daily** with PySpark on Dataproc, optimizing gRPC latency and cloud cost efficiency—infrastructure patterns now applied to AI workloads.
- Designed Delta Lake on GCS with ACID transactions for audit compliance, enabling data versioning critical for ML reproducibility.

PecSmart (IoT & AgTech)

- Built **IoT + ML pipelines** for livestock monitoring with Raspberry Pi and real-time behavioral analysis.

Backend Developer

Jan 2017 – Oct 2017

Fontes Promotora

- Built scalable APIs using **FastAPI, Flask, Django REST** and web scraping at scale.

ENTREPRENEURIAL EXPERIENCE

Founding Engineer & Technical Co-founder

MoveUp AI (San Francisco - Remote)

- Architected **Hybrid RAG with Cognee**: unified vector search (dense) + graph search (symbolic) for complex query routing and entity extraction across unstructured data.
- Built **MCP servers** for multi-tool orchestration across Gmail, Drive, Jira, GitHub with **OAuth 2.0 Gateway** and human-in-the-loop for critical agent actions.
- Developed agents with LangChain/LangGraph, **Mem0 long-term memory**, and model-as-a-judge evaluation for output quality assurance.

AI Engineer & Computer Vision

Trinnix AI Lab

- Agricultural CV systems with **YoloX/Roboflow** (>95% accuracy) + LLM integration (GPT-4o, Claude, Gemini) for intelligent alerts.

TECHNICAL EXPERTISE

AI Platform & LLMs

LangChain, LangGraph, Vertex AI, RAG Evaluation, Prompt Engineering, Tool Calling, Model-as-a-judge, Multi-tool Orchestration.

Retrieval & Knowledge

Hybrid Retrieval (Vector + Graph), Cognee, Semantic Search, Knowledge Graphs, RDF/OWL, Embeddings, Faithfulness/Relevance metrics.

Programming & APIs

Python, TypeScript, Node.js, FastAPI, gRPC, React, SQL, Microservices, REST.

Cloud & Observability

GCP (BigQuery, Vertex AI), AWS, Docker, Kubernetes, Datadog/Grafana, SLOs, CI/CD.

PORTFOLIO & EDUCATION

 pedroreichow.com.br — Interactive Knowledge Graph + Hybrid RAG demo with multi-LLM integration.

B.S. Electrical Engineering | UFSC • **B.S. Systems Analysis** | Estácio