PEDRO REICHOW

pedroreichow3@gmail.com • linkedin.com/in/pedroreichow • pedroreichow.com.br

SUMMARY

AI & Data Engineering leader with 8+ years of experience across fintech, healthcare, and enterprise SaaS. Currently building production AI agent systems at Santodigital (Google Cloud Partner) — shipping 8+ agent deployments with full-stack interfaces, document intelligence pipelines, and dynamic enterprise federation. Built data pipelines processing 500GB+ daily at QI Tech, designed RDF/OWL ontologies and knowledge graphs on Anzo Analytics at Capgemini for fraud detection, and led AI observability across 12+ models at CVS Health under HIPAA/FedRAMP. Co-founded two AI ventures: MoveUp AI (multi-agent orchestration for 15+ enterprise teams) and Trinnix AI Lab (computer vision & knowledge graphs for precision agriculture). Hands-on from PySpark/Spark Streaming at scale to deploying LLM agents with Google ADK, LangGraph, vector databases, and MCP orchestration.

ENTREPRENEURIAL EXPERIENCE

MoveUp AI   San Francisco, CA (Remote)   **Jan 2025 – Jul 2025**

Founding Engineer & Technical Co-founder

- Co-founded AI workforce intelligence platform serving 15+ enterprise teams; designed the entire technical architecture from zero to production deployment.
- Built graph-based multi-agent orchestration system with specialized agents (research, analysis, action) coordinating via shared memory pools and tool-use protocols using LangGraph and Agno.
- Architected knowledge graph infrastructure with FalkorDB and Neo4j, integrating Supabase vector stores and Zep persistent memory for long-term contextual agent reasoning across 50K+ documents.
- Designed and deployed MCP (Model Context Protocol) servers in TypeScript enabling agent-to-agent communication across Slack, CRM, and internal tools with sub-second latency.
- Built end-to-end autonomous workflows with N8N orchestration: agents that research, draft, review, and deliver reports without manual intervention, reducing analyst workload by ~60%.

Trinnix AI Lab   **Feb 2024 – Nov 2024**

Co-founder & Lead AI Engineer

- Co-founded AI consulting lab focused on computer vision, LLM integration, and knowledge graph solutions for agricultural technology and precision farming clients.
- Built and deployed agricultural monitoring systems using YOLOv8 and Roboflow, achieving >95% accuracy in plant disease detection and livestock behavioral analysis across 3 crop/breed types.
- Architected ontological knowledge graphs with Neo4j for agricultural data modeling and business intelligence, enabling semantic queries across geospatial, environmental, and production datasets.
- Integrated multiple commercial LLMs (GPT-4o, Claude 3.5, Gemini 1.5) with advanced prompt engineering chains for automated field reporting, anomaly alerting, and agronomic insights.
- Built scalable full-stack platforms with TypeScript, tRPC, and Node.js serving geospatial analysis dashboards with real-time sensor data overlays and satellite imagery integration.

PROFESSIONAL EXPERIENCE

Santodigital   Google Cloud Premier Partner, Brazil   **Jan 2024 – Present**

Cloud AI Solutions Architect & Agent Developer

- Designed and shipped 8+ production AI agent systems for enterprise clients using Google ADK, building full-stack interfaces (React + FastAPI) that bridge LLM capabilities with business workflows — document analysis, contract review, compliance automation, and customer service bots.
- Built end-to-end document intelligence pipelines: PDF/image ingestion, metadata extraction with structured output parsing, semantic chunking, and vector indexing on Vertex AI + BigQuery, processing 100K+ documents across client deployments.
- Architected dynamic federation integrations connecting AI agents to client ERP, CRM, and internal systems via REST/gRPC adapters, enabling agents to query live data sources and execute actions across heterogeneous enterprise environments.
- Designed production RAG pipelines with offline evaluation suites (RAGAS framework) measuring retrieval precision, faithfulness, and answer relevance; implemented agentic reinforcement loops where observability telemetry feeds prompt versioning, reducing hallucination rates by 35%.
- Built multi-agent orchestration patterns: autonomous compliance monitoring with alert escalation, contract clause extraction agents with human-in-the-loop review gates, and auto-service agents with Zendesk handoff for critical customer interactions.
- Managed training data operations for domain-specific fine-tuning: annotation workflows, data quality validation pipelines, and versioned dataset management on GCS with DVC, supporting model iteration cycles across 5+ client verticals.

CVS Health   **Aug 2024 – Dec 2025**

Senior Data & AI Engineer | AI Observability Lead

- Defined and implemented SLOs for AI pipeline latency (p50/p95/p99) and cost-per-token monitoring using Datadog and Grafana across 12+ healthcare AI models in HIPAA/FedRAMP-compliant environments.
- Built real-time observability dashboards tracking faithfulness, relevance, and groundedness metrics for LLM outputs, enabling prompt version A/B testing that improved accuracy by 22%.
- Led a team of 4 engineers developing enterprise agentic applications with predictive analysis, multi-step reasoning pipelines, and autonomous workflow orchestration using Python and TypeScript.
- Designed data ingestion pipelines processing clinical and pharmacy datasets (200GB+ daily) using Apache Spark and Airflow, feeding ML models for patient outcome prediction.

Capgemini Brasil   **Jan 2022 – Dec 2023**

Senior Solutions Consultant & Knowledge Engineer

- Architected enterprise knowledge graph platform using Anzo Analytics with RDF/OWL ontologies for fraud detection and risk scoring in the financial sector, processing 10M+ entity relationships across distributed semantic stores.
- Designed domain ontologies with Protégé (SKOS, RDF, OWL) and built SPARQL query interfaces on Anzo Analytics, enabling semantic search and complex business intelligence queries across heterogeneous financial datasets.
- Built Python API middleware with FastAPI wrapping SPARQL endpoints, serving as the abstraction layer between the Anzo graph store, relational databases, and front-end applications; handled 2K+ requests/sec

in production.

- Led data modeling initiatives across a 50+ engineer team: defined ontology standards, entity resolution pipelines, and data quality frameworks for financial transaction datasets.
- Implemented ETL pipelines with PySpark and Airflow for daily ingestion of 100GB+ financial records into the knowledge graph, with automated anomaly detection and alerting.

Simulated Reality   Jan 2021 – Dec 2021
Full Stack Engineer | VR, Real-Time & 3D Systems

- Built scalable backend for VR training and healthcare wellness applications using Node.js and React, serving 10K+ concurrent sessions with WebSocket real-time communication and spatial coordinate synchronization for multi-user 3D environments.
- Designed event-driven architecture with RabbitMQ and pub/sub patterns for asynchronous processing of large 3D assets (volumetric video, Unity/Unreal Engine builds, geospatial scene data) totaling 2TB+ weekly throughput across distributed CDN nodes.
- Developed user behavior analytics pipelines capturing spatial interaction data (gaze tracking, hand position, movement patterns), feeding TensorFlow models for adaptive training experiences and personalized wellness program recommendations.
- Engineered blockchain-based asset provenance system for tracking VR content licensing and distribution across partner networks, ensuring IP compliance for medical training modules.

QI Tech   Fintech   Jan 2019 – Dec 2020
Senior Data Engineer

- Processed 500GB+ daily financial transactions using PySpark (Spark SQL, Spark Streaming) on GCP Dataproc clusters, achieving 40% latency reduction through partition optimization and predicate pushdown tuning.
- Architected Delta Lake on GCS with ACID transaction guarantees for regulatory audit compliance; designed data versioning and time-travel capabilities enabling ML experiment reproducibility.
- Built end-to-end ETL/ELT pipelines with Apache Airflow (100+ DAGs) orchestrating data flows from 20+ sources into BigQuery data warehouse and Spark-based feature stores.
- Designed star/snowflake schema data models for the analytics warehouse, supporting real-time dashboards and batch reporting for risk, credit scoring, and regulatory teams.
- Optimized gRPC microservice communication layer between data services, reducing inter-service latency by 55% and cloud compute costs by 30% through resource right-sizing.

PecSmart   IoT & AgTech   Nov 2017 – Dec 2018
Backend Developer & Data Engineer

- Built real-time IoT data pipelines ingesting sensor streams from 500+ Raspberry Pi devices monitoring livestock behavior, processing 50K+ events/hour with Python and Redis Streams.
- Developed computer vision models (OpenCV, early YOLO) for automated animal health assessment, achieving >95% accuracy in disease detection across 3 cattle breeds.
- Designed PostgreSQL time-series data models and batch analytics pipelines feeding predictive models for feed optimization, reducing operational costs by 20% for partner farms.

Fontes Promotora   Fintech   Jan 2017 – Oct 2017
Backend Developer

- Built RESTful APIs using Django REST and Flask, creating a middleware abstraction layer serving front-end applications from multiple database backends (PostgreSQL, MongoDB).
- Developed data collection pipelines using Scrapy for automated financial data extraction from 30+ sources, feeding analytics dashboards and credit assessment models.
- Delivered full-stack solutions with React frontend deployed on AWS (EC2, S3, RDS) with Docker containerization and CI/CD via GitHub Actions.

TECHNICAL EXPERTISE

Languages & Frameworks: Python, TypeScript, JavaScript, SQL, Java (reading proficiency), Scala (reading proficiency) | FastAPI, Django, Flask, React, Node.js, GraphQL, gRPC

AI & ML: LangGraph, LangChain, CrewAI, Agno, Google ADK | RAG Pipelines, Multi-Agent Orchestration, MCP Servers, Prompt Engineering | TensorFlow, PyTorch, OpenCV, YOLOv8 | LLMs (Claude, GPT-4, Gemini, Llama, Qwen)

Data Engineering: Apache Spark (PySpark, Spark SQL, Streaming), Apache Airflow, Apache Flink, Delta Lake, Hadoop (HDFS, MapReduce) | ETL/ELT, Data Modeling (Star/Snowflake), Data Quality Frameworks, Stream Processing

Knowledge & Search: Anzo Analytics, Neo4j, FalkorDB, RDF/OWL/SKOS Ontologies, SPARQL, Protégé | Vector Databases (Supabase, Pinecone, Chroma), Semantic Search, Entity Resolution

Cloud & Infrastructure: GCP (BigQuery, Dataproc, Composer, Vertex AI, GCS), AWS (S3, Redshift, Glue, Lambda, EMR), Azure | Docker, Kubernetes, Terraform | CI/CD, Datadog, Grafana

Specialized Systems: Real-Time 3D/VR (Unity, Unreal Engine, WebSocket), Geospatial Data Processing, IoT Pipelines (Raspberry Pi, Redis Streams), Blockchain Asset Provenance, Distributed CDN Architecture

EDUCATION

Estacio University — B.Sc. Systems Analysis & Development	**2020 - 2023**
Federal University of Santa Catarina (UFSC) — B.Sc. Electrical Engineering	**2017 - 2020**
Federal Institute of Santa Catarina (IFSC) — Technical Diploma, Electrotechnics	**2014 - 2017**