

KEVIN PEREZ-ESPINOZA

AI ENGINEER

New Jersey, United States | kperezespinoza2001@gmail.com

SUMMARY

AI Engineer with hands-on experience building LLM-powered applications, retrieval-augmented generation (RAG) systems, and document intelligence workflows. Experienced with Python, FastAPI, LangChain, LangGraph, OpenAI, AWS Bedrock, and vector search. Combines AI application development with enterprise data governance, workflow automation, metadata quality, and validation.

EXPERIENCE

AI Engineer | Cognizant Technology Solutions — Teaneck, NJ | January 2026–Present

- Develop Python automation for enterprise data workflows, reducing manual validation effort by 70% and improving metadata processing efficiency by 50%.
- Supported migration and validation of 10,000+ metadata assets from Collibra to Alation, achieving 95%+ data accuracy through automated quality assurance.
- Use Python, SQL, and REST APIs to streamline data processing, validate outputs, and identify discrepancies across enterprise datasets.
- Perform data profiling, validation, and issue analysis to improve information reliability and support enterprise AI-readiness initiatives.
- Collaborate with data governance, analytics, and AI teams to improve data accessibility and support enterprise data modernization.

SELECTED AI PROJECTS

SpotiCase | React, TypeScript, three.js, React Three Fiber, Spotify Web API

- Built a Spotify web client around a real-time 3D CD jewel case, with a spring-driven hinge and a disc that spins only while Spotify reports active playback.
- Implemented four playback routes (Web Playback SDK, Spotify Connect remote control, embedded player, and a preview mode) with token refresh and rate-limit handling.

ToneForge AI | C++20, JUCE, Python, FastAPI, Ollama

- Built a guitar-effects plugin (VST3 and standalone) with a C++ DSP library and an eight-slot pedalboard, where a text prompt or song name produces editable pedal settings.
- Built a FastAPI backend that maps a local LLM's tone description to pedal parameters, with a deterministic fallback when the model is unavailable.

LegalRAG | Python, FastAPI, AWS Bedrock, Claude, ChromaDB Cloud, Sentence Transformers

- Built a legal-document RAG platform for grounded question answering over specialized documents.
- Integrated AWS Bedrock with Claude and vector-based retrieval, using sentence-transformer embeddings and ChromaDB Cloud for semantic search.

OnTract Contract Management System | OpenAI, Supabase, LangSmith

- Collaborated on a team-developed AI-powered contract management platform for LLM-assisted document analysis and contract workflows.
- Contributed to an application integrating OpenAI with Supabase-backed services and LangSmith for LLM tracing and observability.

TECHNICAL SKILLS

AI & Agentic Systems: LLM Applications, RAG, LangChain, LangGraph, OpenAI API, AWS Bedrock, Semantic Search, Vector Embeddings, Prompt-Based Workflows

Programming & Backend: Python, SQL, Java, JavaScript, TypeScript, C++, FastAPI, REST APIs

Data & AI Infrastructure: ChromaDB, Supabase, Amazon S3, Pandas, NumPy, Sentence Transformers, LangSmith

Enterprise Data & Quality: Data Profiling, Data Validation, Metadata Management, Data Quality, Data Lineage, Collibra, Alation

Tools: Git, GitHub

EDUCATION

Rider University — Bachelor of Science in Computer Science | 2024