

Portfolio AI assistant

Jordi Hako Nintcheu · 2026-07-14

A RAG chatbot embedded on this site, answers questions from the real project write-ups instead of generic guesses.

Problem

Most portfolio sites are static, a recruiter has to click through every project to find what they need. This one answers directly instead: a chat widget on the site itself, grounded only in the actual case studies. It also doubles as the clearest proof of skill on the whole site, since it doesn't describe a RAG pipeline, it runs one, live.

Approach

Python and FastAPI on the backend, Pinecone for the vector store, OpenAI for both embeddings and generation. Deployed as its own Vercel project rather than folded into this site's deployment, so the chat widget calls it cross-origin instead of sharing a codebase. Content gets chunked from the site's own MDX case studies and notes, embedded, and indexed, then retrieval pulls the closest chunks for a given question and generation answers from them alone.

The one real design correction: filtering weak retrieval matches by a flat similarity score threshold didn't work, OpenAI's embedding models compress similarity into a narrow band regardless of actual relevance, so there's no clean line to threshold against. Replaced it with structured outputs, the model reports which retrieved chunks it actually used to answer, and only those get cited as sources.

Result

Live on this site right now, bottom-right corner. Indexes the project and note content, answers grounded and cited correctly, no unrelated projects showing up as sources for an on-topic question. Stateless (no memory across questions), no streaming, and reindexing after new content ships is still a manual step, not automatic yet.

What I'd do differently

Would have logged real retrieval scores before guessing at a similarity threshold, instead of building and testing an approach that turned out to be the wrong tool for the problem.

Part of an ongoing series of applied AI and engineering projects. Full write-up and live links at jordihako.com.