

Antoine Rebeiz

AI Systems Engineer (LLMs, RAG, Data Pipelines, AWS)

✉ @rebeizantoine@gmail.com 🌐 <https://antoineportfolio.site> <https://www.linkedin.com/in/antoine-rebeiz> ☎ +961 76045061

👤 Professional Summary

AI Systems Engineer with 2+ years of experience designing and deploying LLM-powered applications and data pipelines. Specialized in retrieval-augmented generation (RAG), multi-LLM orchestration, and AI workflow automation. Experienced in semantic search, embeddings, and structured output generation, with a strong backend foundation (Node.js). Focused on building reliable, scalable AI systems and improving output quality through prompt design and evaluation.

📁 Work History

Nexus IT Solutions UK-Yorkshire(remote) | Aug 2024 - Feb 2025

Full-Stack Web Developer

- Built and maintained MERN stack applications, improving performance, scalability, and UI/UX consistency
- Collaborated via GitHub (PRs, code reviews), accelerating development cycles and reducing integration issues
- Integrated third-party APIs and AWS services to support high-volume requests and enhance platform functionality

Freelance Full-Stack Engineer (SaaS & E-commerce) | Mar 2024 - Jun 2025

Full-Stack & AI Developer

- Promptsphere: Engineered a multi-LLM system with semantic retrieval, structured output generation, and prompt routing
- DocuMind: Implemented pipeline: document ingestion → chunking → embeddings → semantic search → LLM answering, Using vector-based retrieval to improve response relevance
- AI Workflows (n8n): Built pipelines for classification, enrichment, and automated routing, Enforcing structured JSON outputs for downstream processing

★ Skills

- AI & Data Systems:** LLMs (OpenAI, OpenRouter), RAG, Embeddings, Semantic Search, Prompt Engineering, AI Output Evaluation
- Backend & Data Engineering:** Node.js, Express.js, REST APIs, Async Processing, Queues (BullMQ), Data Pipelines
- Cloud & Infrastructure:** AWS (Certified), Scalable API Design
- Frontend:** React.js, Next.js, TypeScript

🎓 Education

Civil Engineering

Conservatoire National des Arts et Métiers
2018 - 2025
Master's Civil Engineering

Full Stack Web Development

AUF | Simplon.co
2023 - 2024
Completed an intensive 6-month (720-hour) program focused on the MERN stack and production-ready web applications, including e-commerce solutions.

📜 Certifications

- AWS Certified Cloud Practitioner - Amazon Web Services Training and Certification Issued Jan 2025 - Expires Jan 2029
- AWS Certified AI Practitioner - Amazon Web Services Training and Certification Issued Sep 2025 - Expires Sep 2029