



BENOIT GAUDIERI

AI Engineer | LLM Systems | Full-Stack

Naples, Italy • +39 339 561 4592 • ben.gau1987@gmail.com

[LinkedIn](#) • [GitHub](#)

SUMMARY

Technical Lead with 6+ years in production web engineering and a last-two-year focus on LLM systems. Architected the headless migration of Forbes Italia (one of Italy's highest-traffic editorial sites) integrating AI-powered editorial features; ship RAG pipelines, multi-LLM orchestration (OpenAI, Claude, Gemini) and agent workflows with LangChain and LangGraph. Come from comparative literature before engineering — which shapes how I design prompts, conversational UX, and user-facing AI surfaces. Native Italian, bilingual in English.

SELECTED AI / LLM PROJECTS

- **Forbes Italia — Editorial AI platform (2024-present).** Architected the headless WordPress → Next.js + Laravel + Python migration and integrated AI-assisted editorial features (summarization, tagging, semantic search) on a site handling high daily traffic.
- **RAG systems for context-aware assistants.** Designed and shipped retrieval-augmented pipelines on Supabase/PostgreSQL vector stores, with chunking, hybrid retrieval and prompt orchestration tuned for conversational accuracy in enterprise contexts.
- **Multi-LLM orchestration with LangChain / LangGraph.** Production agent workflows spanning OpenAI GPT, Anthropic Claude and Google Gemini, with routing, fallback and prompt-engineering patterns that keep cost predictable and quality stable.

EXPERIENCE

Technical Lead / Full-Stack Developer

[IDI Group S.R.L.](#)

Apr 2024 – Present

Naples, Italy

- Lead the AI integration practice: design RAG pipelines, vector-DB schemas and LLM orchestration layers used across multiple enterprise clients.
- Architected Forbes Italia's headless migration (Next.js + Laravel + Wordpress) with embedded AI editorial features — one of the firm's flagship 2024-2025 deliveries.
- Drive prompt engineering standards and evaluation practices across the team; own model selection, cost monitoring and reliability for production LLM features.
- Mentor engineers on AI integration patterns and own architecture decisions alongside stakeholder coordination and delivery quality.

Full-Stack Web Developer

[The Medix](#)

Feb 2023 – Apr 2024

Naples, Italy

- Led frontend for healthcare-management products (Angular/React) and a React Native patient-engagement app, focused on usability and performance.
- Led development for the U.S. Salernitana 1919 (Serie A) project: set up the Laravel architecture from scratch and coordinated the delivery team.
- Orchestrated L'Espresso editorial migration: high-traffic legacy blog → headless CMS with Next.js frontend, including XML data transformation and content integration.
- Managed a 3-person frontend team, coordinating delivery workflows and code-quality standards.

Junior Web Developer

LAB46

Nov 2020 – Feb 2023

Naples, Italy

- Angular frontends and custom vTiger/PHP CRM extensions in a fast-paced digital-agency environment.
- Introduced Jira and Agile workflows that improved sprint velocity and team coordination.
- Managed Debian server infrastructure for multiple client sites; built high-performance landing pages with Vite/Parcel and GSAP animations.

SKILLS

AI & LLM: LangChain, LangGraph, OpenAI GPT, Anthropic Claude, Google Gemini, Ollama, MCP servers, RAG pipelines, vector databases (pgvector/Supabase), prompt engineering, Stable Diffusion

Backend: Node.js, Express, Python, Laravel, PHP

Frontend: React, Next.js, Angular, Astro, Vue, Svelte, Tailwind CSS, Three.js, GSAP

Mobile: React Native, Flutter

Data & Infra: PostgreSQL, MySQL, MongoDB, Supabase, Firebase, Docker, Linux/Debian

Testing: Pytest, Flasgger, unit testing

Languages: English (native), Italian (native), French (intermediate), Japanese (basic)

EDUCATION

- **Full-Stack Web Developer Bootcamp** — Boolean Careers, Milan (2020)
- **iOS Foundation Program** — Apple Developer Academy, Università L'Orientale, Naples (2017)
- **Comparative Literature & Languages (coursework)** — Università L'Orientale, Naples