

Omar G. Nagy

AI Product Engineer

Egypt (remote, GMT+2) · omar@neurасcale.org · [+20 150 047 7334](tel:+201500477334)
omargnagy.com · github.com/omarnagy91 · linkedin.com/in/omargnagy

SUMMARY

AI Product Engineer with 9 years shipping production software, the last 3+ building LLM systems that run in production. Founder-operator of NeuraScale: products in production and 2 open-source tools, built and operated solo, including a multi-tenant POS SaaS and a medical exam-prep platform with 10,700+ questions. Cut LLM costs 28% with measured, published methods. Full EU timezone overlap, 4+ hours with US East.

SKILLS

AI / LLM	Claude API (Anthropic), OpenAI API, multi-agent orchestration, agent tool use / function calling, RAG, vector search (pgvector, Qdrant), prompt engineering, prompt caching, eval harnesses, LLM cost optimization, Model Context Protocol (MCP)
Languages	TypeScript, JavaScript, Python, SQL, Bash
Web	Next.js (App Router, RSC, Server Actions), React, Node.js, FastAPI, REST APIs, Tailwind CSS
Data / Infra	PostgreSQL (RLS, migrations, query tuning), Supabase, Vercel, GitHub Actions, CI/CD, Sentry, GA4
Testing	Playwright, Cypress, end-to-end coverage on critical paths

EXPERIENCE

NeuraScale · Founder & AI Product Engineer Feb 2024 – Present · Remote (Egypt)

- Built and operate production products solo (RetailOS multi-tenant POS SaaS, MedPrüf exam prep, Bridge Sourcing B2B, Harmonia POS, Blink cafe OS) on TypeScript, Next.js, Supabase/PostgreSQL, and Vercel, serving live paying users.
- Shipped MedPrüf, an Austrian medical-licensing exam prep platform, growing it to 10,700+ active questions across 3 exam types and 16 specialties, by building an AI content pipeline with strict-format explanation generation in German plus 7 locales, subscription payments, and a 7-tab admin dashboard.
- Cut portfolio-wide LLM spend 28% with no quality regression by building model routing, prompt caching, and context-budget enforcement; published the per-change breakdown publicly.
- Took RetailOS from hand-managed to invite-only multi-tenant SaaS in a 48-hour sprint by shipping 12 PRs and 12 PostgreSQL migrations (RLS hardening, webhook signature auth, idempotency, server-side re-pricing) with zero data loss.
- Raised extraction accuracy from 82% to 96% on Bridge Sourcing’s supplier-data pipeline (20 structured fields per record) using hand-labeled calibration sets, deterministic validation outside the model, and human-review acceptance gates; published the full engineering breakdown.
- Scaled a one-person engineering org by designing a multi-agent build harness (research, build, review, and test lanes with human approval gates); open-sourced the agent memory layer (Mnemonic) and the eval tooling (llm-eval-ci).
- Selected as a paid evaluator of agentic coding assistants (primarily Claude Code), reviewing and validating AI-generated changes across advanced production codebases (under NDA).

Independent (Consulting) • AI & Automation Engineer

Sep 2022 – Jan 2024 • Remote

- Reduced an enterprise client's LLM spend by building cost guardrails (daily caps, cheap-prompt-first routing, context budgets) with no measured quality regression; the pattern became standard across all my later products.
- Built production web-data extraction pipelines (retry, deduplication, fallback chains) and lead-scoring microservices for SME clients pivoting into AI integration.
- Shipped a production multi-agent pipeline in early 2023 (RAG retrieval plus tool-calling orchestration with hand-rolled JSON-mode prompting and retries) before mainstream agent frameworks existed.

Independent • E-Commerce Engineering Consultant

Mar 2021 – Aug 2022 • Egypt / Remote

- Drove ~35% conversion lift for seller cohorts on Taager, a \$6.4M-funded social-commerce platform scaling past 100+ active sellers, through funnel instrumentation and seller-side conversion work.
- Replaced spreadsheet stock counts with a live inventory-sync engine across 4 warehouses for a textile manufacturer, including a QR-code loyalty MVP, delivered in 6 weeks.
- Migrated 3 Shopify storefronts to custom Next.js with Stripe checkout, lifting Lighthouse scores from 38 to 92.

mBalloon • Co-Founder

Feb 2020 – Feb 2021 • Alexandria, Egypt / Remote

- Onboarded 15+ MENA SMBs onto Shopify with paid acquisition during the COVID e-commerce boom; 3 niche stores sustained 2.5-3.2x ROAS for 6+ months.
- Cut per-client setup time from 3 weeks to 5 days by building a reusable Shopify theme and funnel template.

Freelance • Front-End Developer

2019 – 2020 • Remote

- Migrated a 60K-LOC jQuery codebase to React with Cypress E2E coverage on critical paths, cutting time-to-interactive from 3.5s to 2.1s with zero production breakage (additive migration, feature flags).

LibreFit • Founder

2017 – 2018 • Egypt

- Built a smartphone computer-vision body-measurement product end to end (Xamarin client, OpenCV pipeline, server-side calibration) reaching $\pm 1-2$ cm accuracy; sold 2 contracts to MENA tailoring shops.

OPEN SOURCE & WRITING

- **Mnemonic:** Self-hosted agent-memory layer (Qdrant vector storage, retrieval API). github.com/omarnagy91/mnemonic
- **llm-eval-ci:** CI-integrated eval harness for LLM outputs. github.com/omarnagy91/llm-eval-ci
- Writing with real production numbers at omarnagy.com/writing, including:
 - [How I cut our LLM bill 28% without changing models](#)
 - [Bridge Sourcing: how I moved scrape accuracy from 82% to 96%](#)

EDUCATION

BSc, Computers & Informatics • Suez Canal University, Ismailia • 2014 – 2018 • First Class Honours

LANGUAGES

English (fluent) • Arabic (native) • German (A2, learning)