

Khalfoun Mohamed El Mehdi

AI / Agent Engineer · Full-Stack Engineer

Algiers, Algeria · Open to remote · contact.khalfoun@gmail.com · +213 556 807 712

linkedin.com/in/usecallback · github.com/useCallback

SUMMARY

AI/agent engineer and full-stack engineer with six-plus years building production software. Over the past two years I have shipped three production AI agent systems: a sales agent that answers product questions and closes cash-on-delivery orders in chat, an AI receptionist that qualifies leads and books appointments across Instagram and WhatsApp, and a document-intelligence pipeline that turns receipts into structured, tax-compliant records. I build the system around the model: tool-calling and orchestration, retrieval, evaluation and continuous improvement, and the integration work that makes an agent usable inside a real product.

Core AI agents & tool-calling | LangGraph | LangChain | RAG & vector search | evaluation harness engineering | prompt engineering | TypeScript | Python | Node.js | PostgreSQL | GCP

EXPERIENCE

The 2024–2026 engagements below ran concurrently, alongside Algorithme Agency.

Kommerce — Co-Founder & Engineer

2025 – Present

Multi-tenant, cash-on-delivery-first e-commerce SaaS for the Algerian/MENA market. Merchants build a storefront and run orders out of it. Customers pay cash at the door, so every order is confirmed before it ships — by the AI sales agent in chat or by a person phoning from the built-in call-center CRM. Co-founded with a partner; in production.

- **AI sales agent:** one agent core across storefront chat, WhatsApp and handoff to the call centre. It answers product questions from the merchant's catalogue using retrieval, replies in Algerian darija, and completes a cash-on-delivery order inside the conversation.
- **Order confirmation flow:** the agent collects and confirms the order in chat, with a deterministic check on the final confirmation step and anything ambiguous handed to a person in the call-center CRM rather than resolved by the model.
- **Agent quality loop:** agent changes are regression-tested against recorded conversations before shipping, so a change cannot quietly degrade how the agent handles a real case.
- **Storefront and page builder:** a WYSIWYG editor for assembling store pages block by block, with undo/redo, live preview and right-to-left Arabic-first theming — and an AI assistant on the same surface that creates pages, edits blocks and restyles a store from a prompt.
- **Multi-tenant platform:** the backend behind storefronts, orders and merchant accounts, with roughly 12 courier APIs behind one integration layer that routes each order to the carrier serving its commune.

Algorithme Agency — Founder — web development and AI agency, Algiers

Sep 2024 – Present

The vehicle for the concurrent client and product engagements listed here.

Velya — Lead Full-Stack Engineer

Sep 2025 – Jun 2026

AI-receptionist SaaS for clinics and service businesses: turns an Instagram/WhatsApp inbox into an agent that answers from the business's own material, qualifies leads, books appointments and hands over to staff. Lead engineer, roughly 80% of the build.

- **Conversational agent:** tool-calling for appointment booking, lead capture and knowledge-base answers across Instagram and WhatsApp, including understanding the images, reels and stories customers send.
- **Human-like conversation:** message debouncing, so a customer typing in three quick bursts gets one considered reply rather than three; read and seen states so the conversation behaves the way people expect; and understanding and replying in an Algerian tone rather than formal translated Arabic.
- **Continuous evaluation and improvement:** conversation quality is scored automatically, prompt and model changes are A/B tested against it on a from-scratch statistics library, and the results feed suggested prompt and setting changes back to the operator.
- **Booking and billing backend:** the services behind the agent — calendar and appointment handling, lead records, per-organisation configuration, and usage-metered subscription billing across two payment providers.

Wateer — AI Engineer

Oct 2025 – Mar 2026

OCR-AI: LLM document-intelligence platform for a Saudi point-of-sale provider, turning Arabic and English receipts and invoices into a validated, structured record shaped for Saudi tax compliance (ZATCA).

- **Document extraction agent:** a self-correcting extraction workflow that reads a receipt or invoice, checks its own output against the expected shape and retries the fields it got wrong — degrading gracefully on a poor scan rather than failing the whole document.
- **Per-field confidence:** a multi-model ensemble whose outputs are compared against one another to score confidence field by field, so downstream systems know which values to trust and which to route for human review, without a model grading its own work.
- **Webhooks and batch processing:** real-time extraction delivered over webhooks for single-document flows, plus a batch pipeline for bulk back-processing of large document sets.

- **Long-document handling:** chunking that keeps multi-page invoices inside the model's context budget, and an accuracy harness used to measure extraction quality as the pipeline changed.

Osen Express — Full-Stack Engineer, freelance

2024 – 2026

Delivery and fulfilment platform for the Algerian market, covering an order from merchant stock through to a paid-out merchant: warehousing, dispatch, last-mile delivery, returns, and the cash collected on delivery. Freelance via a partner agency.

- **Fulfilment and inventory:** merchant stock held in warehouses and shipped from it on order, with stock levels, picking and returns handled as part of the same order lifecycle.
- **Dispatch and last mile:** orders routed across hubs and out to the customer's door, with driver dispatch, live tracking, a pricing engine, and one customer-facing status derived from the pickup, transfer, delivery and return states underneath.
- **Cash-on-delivery settlement:** an append-only ledger tracking cash custody from driver to facility to merchant payout, each transfer a new immutable entry rather than a balance update, with reconciliation and batch remittances.
- **Merchant integrations:** partner APIs and webhooks so merchants keep running the COD management tools they already use — EcoManager, EcoTrack, COD Pilot — against this platform.

3M Express — Backend Engineer — Algiers, Algeria

Jun 2023 – Sep 2024

- **Routing engine:** owned the routing domain — pathfinding and vehicle routing that decides how an order moves between facilities and out for delivery, across **300,000+ parcels**.
- **Middle mile and facilities:** transfer between hubs and the facility model underneath it — what each holds, what it owes onward, how an order is handed between them — plus the order fulfilment flows on top.
- **Internal tooling:** the internal admin platform, the release-qualification integration-test suite, and the error-logging and monitoring frameworks.

Riverdone — Full-Stack Engineer, London, UK (remote)

Mar 2022 – May 2023

- Led web and mobile development for client products including Finimize, serving 100,000+ active users.

TemtemOne — Full-Stack Engineer, Algiers, Algeria

Nov 2021 – Mar 2022

- Built the back-office and administration dashboards in React, Node.js, TypeScript and MongoDB, for a platform now past **1M+ downloads**.

Techsio — Lead Frontend Engineer, Algiers, Algeria

Sep 2020 – Dec 2021

- Owned the customer-facing frontend of an e-product sales platform across North Africa, the Middle East and Europe as lead frontend developer, serving 60,000+ active users.

Iwello — Full-Stack Engineer, Lagos, Nigeria (remote)

Oct 2019 – Aug 2020

- Contributed telemedicine features for online appointments and teleconsultation, including real-time chat over Socket.IO.

SKILLS

AI & agents AI agents & tool-calling | LangGraph | LangChain | agent orchestration | agent memory & context management | RAG & vector search | multimodal (vision) | multi-model ensembles | conversational UX (debouncing, presence, dialect) | OpenRouter | LangSmith

Evaluation & improvement Evaluation harness engineering | golden-case & regression suites | LLM-as-judge with deterministic scorers | A/B experimentation | prompt engineering & versioning

Backend & data TypeScript | Python | Go | Node.js | NestJS | FastAPI | PostgreSQL | pgvector | Redis | MongoDB | Kafka | multi-tenant SaaS | event-driven systems

Infrastructure & frontend Docker | GCP (Cloud Run, Cloud SQL, Pub/Sub) | Terraform | AWS | CI/CD | React | Next.js | React Native | Stripe | WhatsApp & Instagram APIs

OPEN SOURCE

- Fixed a push-subscription configuration bug in the Google Cloud Pub/Sub Node.js SDK (1M+ weekly downloads); merged documentation fixes into Apple's m1-dx and Microsoft's UniLM.

EDUCATION

University of Algiers — Medicine, M.D., studied alongside full-time engineering work

Sep 2019 – Present

Certifications: EFSET English — 82/100 (Proficient) | Convolutional Neural Networks in TensorFlow, DeepLearning.AI