

JHÉSSINY MATTOS

AI Engineer · Agents · RAG · Google ADK · 5 years software engineering depth

Brazil (UTC-3) · Open to US & EU · jhessinymattos@gmail.com · linkedin.com/in/jhessiny-mattos · jhessiny.me

SUMMARY

Software Engineer with 5 years of experience— now centered on AI engineering: Claude API, RAG pipelines, agents, and MCPs. I build AI products where humans calibrate intent, and AI executes volume. My engineering background means I integrate AI at the system level — not as a feature on top, but as a core architectural layer. Experienced across full-stack B2B and B2C SaaS with cross-functional remote teams.

PROFESSIONAL EXPERIENCE

Software Engineer — AI · Softo — Remote-first software consultancy (~50 engineers) · Feb 2022 – Present

Software Engineer on AI-integrated SaaS for international clients. Five years across React + TypeScript, Node.js, and PostgreSQL — grew into fullstack on the current product, an AI conversation-analysis dashboard in production on NestJS + AWS.

Responsibilities

- Fullstack on Revla, a production AI system for conversation analysis (NestJS + React on AWS) — owned frontend architecture and contributed across the backend and infrastructure layers.
- Implemented LLM output guardrails on Revla: prompt sanitization, response filtering, output length limiting, and fallback handling for unexpected model responses.
- Instrumented Revla with OpenTelemetry and event capture to database for observability across the AI pipeline; used Clarity and New Relic on other products.
- Frontend lead on CollaMap, a multi-tenant scientific knowledge platform (3+ years) — interactive canvas, custom rich-text editor, multi-tenant auth, role-based access.
- Maintained a company-wide React + TypeScript design system adopted across 8+ products.
- PR review, mentoring, technical interviews, hiring decisions.

Impact

- Designed and shipped a production LLM system on Claude API — multi-stage novelty detection (keyword pre-filter → Claude classification) over PubMed at \$0.17/run, 66% cost reduction. Full-stack POC in 3 days. Presented architecture and cost model to the client.
- Built CSV conversation import end-to-end (upload UI → processing pipeline → email notification) — feature that secured the first signed contract and handles 50k+ conversations/day in production.
- Drove the Unleash production rollout end-to-end — caught a v7.6.3 UI regression, diagnosed an Aurora SSL boot failure.
- Reduced TTI by 60% on data-heavy interfaces via code splitting, virtualization, and memoization.

Stack: Claude API, RAG, pgvector, MCPs, Claude Code, React, TypeScript, Node.js, NestJS, Python, PostgreSQL, AWS, OpenTelemetry, Unleash, Docker, Redux, Zustand, Tailwind, Jest, Cypress, Vitest.

Software Engineer · Munddi — B2B2C retail marketplace · Apr 2021 – Jun 2021

- Built React + TypeScript screens for the consumer app: product feed, store-locator map, brand profile, recommendation surface.
- Integrated location and recommendation APIs; implemented state management and responsive layouts across consumer flows.
- Migrated legacy jQuery codebase to React + TypeScript with Material UI and Formik validation.

Software Engineer — Open Banking & Pix infrastructure · Ame Digital / Bit Capital · Jul 2021 – Feb 2022

- Engineering role on Brazil's Open Banking and Pix rollout at Central Bank scale (millions of users, millions in transaction value).
- Automated transaction reprocessing in TypeScript; built Metabase dashboards for transaction health and compliance signals.
- Monitored production services in Datadog — alerts, dashboards, and traces for a high-throughput payments environment.
- Translated Central Bank regulatory documentation into engineering specs; contributed to transaction-rules decisions.

Software Engineer · Freelance · Oct 2020 – Jan 2022

- Built full-stack web apps for small businesses end-to-end: React + TypeScript frontend, Node.js + PostgreSQL backend, scoping → deploy.

PERSONAL PROJECTS

Vegajus — Legal-tech SaaS for solo lawyers and small firms — shipping to production

- Lead engineer and product owner. My legal background drives the product; I lead frontend and co-build the backend.
- Core: Meta Ads management + WhatsApp automation + CRM with conversation-to-case conversion via Claude API.
- Stack: React, TypeScript, Java, PostgreSQL, Supabase, Claude API, WhatsApp Business API.

StudyFlow — *Visual concept-map study platform with AI surfaces*

- React Flow canvas with BYOK demo; AI surfaces: chat onboarding, inline writing feedback, YouTube → concept ingest.
- RAG pipeline with pgvector for semantic retrieval across ingested content.
- Stack: React, TypeScript, Python, React Flow, Node.js, pgvector, Claude API, Supabase.

SKILLS

AI/LLM: Claude API, Anthropic SDK, RAG pipelines, pgvector, embeddings, agents, MCPs, Claude Code, prompt engineering, structured output, guardrails, output validation

Frontend: React, TypeScript, Next.js, Redux, Zustand, TanStack Query, Tailwind, shadcn/ui, MUI

Testing: Jest, RTL, Cypress, Vitest

Backend & Infra: Node.js, NestJS, PostgreSQL, Python, Java, FastAPI, Docker, AWS Lambda, Supabase

Observability: OpenTelemetry, New Relic, Datadog, Microsoft Clarity, Metabase

LANGUAGES

Portuguese — Native · English — Fluent · French — Advanced