

JAMES LORENZ SANTOS

AI ENGINEER | AGENTIC SYSTEMS, AI-FIRST FULL STACK DEVELOPMENT AND AUTOMATION

+63 936 081 1399 | jameslorenzsanatos@gmail.com | linkedin.com/in/james-lorenz-santos-720776251
github.com/jamessuuu | agentjames.vercel.app | Quezon City, Philippines | Open to remote roles

SUMMARY

AI engineer who ships AI features to real users and keeps them working after launch. I build agentic systems that do real work, including multi agent pipelines, retrieval grounded assistants, and typed tool surfaces such as MCP and OpenAPI that let AI query and act on production systems safely. I also build the evaluation harnesses and verification gates that make model output reliable enough to put in front of a paying customer. TypeScript end to end across Next.js front ends and Node and NestJS service APIs, with Python where the work calls for it. Delivered across managed IT services, legal marketing, fintech, healthcare operations, dental, apparel and construction, usually as the only AI engineer on the team and often as the whole delivery function. Eight Anthropic and PMI AI credentials earned between January and June 2026. Summa Cum Laude and Rank 1 in the IT Department, University of Santo Tomas.

WHAT I BRING

- **Agentic systems in production, not demos.** Multi agent pipelines with model routing per task, agents that read live business data, classify and score it, and return a recommended next action with a drafted artefact. Structured output is schema validated at every step, and agents propose writes rather than executing them, so a human stays accountable.
- **Assistants that take actions safely.** Typed, scope gated tool surfaces over real systems using MCP and OpenAPI 3.1 generated from Zod schemas, with per agent API keys, role based access and every agent action attributed in an audit trail.
- **Reliability engineering for model output.** Deterministic verification gates, golden task regression suites that fail the build on regression, a named failure taxonomy covering hallucination, over refusal, malformed tool calls and cost blowups, and hard loop caps that escalate to a human instead of letting an agent spiral.
- **AI-first delivery speed at production quality.** An agent plugin in two days, an internal operations platform in two days, a full stack AI credit application in 30 days, and a site generation pipeline that cut delivery from weeks to hours, all with test suites and CI behind them.
- **Automation that removes real cost.** Triage agents, connectors and integration pipelines that cut service delivery time by 30%, and AI native rebuilds that cut manual operational work by 80%.

PROFESSIONAL EXPERIENCE

AI and Automation Developer, Contract | Aldridge, Managed IT Services, United States (Remote)

May 2026

- Pioneered the AI practice across internal operations and client service delivery, building triage agents, connectors and automation pipelines that **cut service delivery time by 30% within three months**.
- Built AI workflows, internal tools and copilots using **Microsoft Copilot and Copilot Studio** alongside custom agents on the Anthropic Claude API and Model Context Protocol, so engineers could query and act on live operational data instead of hunting through consoles.
- Integrated AI across the MSP toolchain, covering **RMM platforms including ConnectWise Automate and PSA systems including ConnectWise Manage, Neo and Thread**, connecting ticketing, monitoring and client data into a single automated workflow.
- Shipped automation applications, data pipelines and internal web apps, and led AI enablement across the organisation with reusable agent patterns and onboarding so non specialists could use AI safely in daily work.

AI Automation Developer, Contract | Lift Legal Marketing Pty Ltd, Sydney, Australia (Remote)

Jan 2026 - Present

- Designed and shipped **Law Firm Builder Pro**, an agentic plugin and multi agent platform that generates complete production websites from a brief. **20 or more specialised agents across a six phase pipeline** covering strategy, design, build, verify and deliver, with model routing per task so judgment heavy agents run on frontier models and high volume agents run on faster tiers. Cut delivery from weeks of manual build to **2 to 4 hours per site**, and it now underpins **20 or more client platforms**.
- Made that output trustworthy with **124 automated verification gates** and a three rule honesty protocol covering gate output attestation, read only tooling policy and mandatory escalation instead of gate editing, producing verified accessibility 100 and PageSpeed 95 to 96 across 274 live URLs.
- Own the AI layer across the client fleet, including content generation and analysis, analytics grounded content scoring, AI insight surfaces, and a site wide RAG assistant whose answers are personalised per page and use case through a multi stage conditional retrieval pipeline.
- Built a four agent pipeline over live CRM data that classifies each account's funnel stage and engagement tier, then returns a next best action with a personalised outreach draft, proposing CRM writes for human approval rather than executing them.

- Built an analytics and content intelligence SaaS on **TypeScript and Convex** with headless GA4, Search Console and Tag Manager integration. LLM optimisation output passes through a field whitelist and schema validator before persistence, and every merge is gated by a **417 test CI suite**.
- Owned the full lifecycle as sole engineer on most builds, covering requirements, architecture, UI and UX, QA and CI/CD on GitLab and GitHub, sustaining **99.5% uptime** with zero downtime releases.

Software Engineer | Kredit Hero, Inc., BGC, Taguig, Philippines

Jul 2025 - Dec 2025

- Pioneered the company's AI integrations using **Google Vertex AI** for document processing and natural language processing, plus text and chat agents embedded directly in the credit application flow.
- Built a **Custom AI Processor** feature that let non technical users compose their own processors and processing flows natively in the product, so each client could encode its own business rules for evaluating applications. Platform adoption rose **50%**.
- Delivered a full stack AI powered credit application from concept to production in **30 days** using Next.js, Python and Google Cloud Platform, launched to **1,000 or more users**.
- Ran three concurrent projects as hybrid project manager, business analyst and QA, hitting every milestone **15% ahead of schedule**, and learned an unfamiliar GCP stack in two weeks while raising team output velocity by **50%**.

Full Stack Developer, Internship to Engineer | eBiZolution, Inc., Makati, Philippines

Jan 2025 - Jun 2025

- Co led the redesign of the company website and a reusable admin template, shipping the production application in Next.js, TailwindCSS and Material UI. Received 95% positive feedback from business unit heads and clients, and the highest performance evaluation of five interns.

INDEPENDENT AI CONSULTING AND FREELANCE ENGAGEMENTS

AI Engineer and Consultant | Kelowna Dental Centre, Canada (Remote)

2026

- Led the practice's transition to **AI native operations**, building a **marketing agentic plugin** that became the foundation layer of their platform, plus a **central control platform** unifying marketing and business operations into one system of record.
- Owned AI era search and paid acquisition across SEO, AEO and GEO alongside Google and Meta Ads, improving process efficiency and return on marketing spend.

Full Stack Architect and Developer | One Nadela, healthcare operations group, Philippines

2026

- Designed and built an internal operations platform covering inventory, tasks, recurring workflows, dashboards and reporting as an installable offline capable PWA, **built in two days** using AI assisted development.
- Made the platform **agent operable** with 24 scope gated MCP tools plus REST and OpenAPI 3.1 generated from Zod schemas, secured by per agent API keys, role based access and full attribution in the activity feed.
- TypeScript end to end on Cloudflare Workers, Hono, D1 and Drizzle behind React, Vite, Tailwind and shadcn/ui, with **289 tests** running in about three seconds against real migrations.

AI Feature Consultant | Apparel technology startup, New York, United States (Remote)

2026

- Designed and built the AI layer of a consumer clothing application, letting users generate original designs from a prompt, refine an existing design iteratively, and find visually similar designs. Integrated **Google Gemini and OpenAI** APIs behind a single provider agnostic interface, owning prompt design, image pipelines and output quality control.

AI Solutions Consultant | Commercial construction contractor, United States (Remote)

2026

- Built structured **Claude Projects** that turned ad hoc bid preparation into a repeatable pipeline covering document intake, scope extraction and response drafting, then extended the same pattern to adjacent construction workflows.

SELECTED ENGINEERING AND OPEN SOURCE

- **Grounded AI Insight Engine** (clubscope-insight-engine.vercel.app, source on GitHub). A full stack analytics and assistant platform on **TypeScript, NestJS and Next.js**, built on one constraint: the language model never produces a number. It selects typed analysis tools, every figure it writes must cite its evidence, and a deterministic verifier recomputes each figure from source before anything renders, failing closed when it cannot. Includes explainable non model risk scoring, an assistant that proposes actions for human confirmation with an audit trail, and a live reliability panel that demonstrates a fabricated figure being blocked. **287 tests, CI green, eval suite gating the build**.
- **Personal agentic engineering ecosystem**. 44 specialised agents and 48 skills across eight divisions, on a deterministic enforcement harness featuring verifier gated bounded loops where the runner rather than the agent executes the verifier, so an agent cannot fake a green result. Includes a golden task regression runner, boundary and secret leak guards, and a 28 case self test on the guards themselves. This is what lets one engineer hold several production workstreams at once without dropping quality.

- **15 public reference implementations** at github.com/jamessuuu, all CI green, deployed and HTTP verified, with **1,043 or more aggregate tests**. Themes include retry and idempotency reliability, MCP protocol testing with SSRF guards, and eval harnesses proven byte identical across runs.
- **agentjames.vercel.app**, a Next.js 16, React 19 and TypeScript strict portfolio console with 233 tests including a 20,000 pair fuzz suite. **joinklik.app**, a live two sided matching platform with passwordless auth, server enforced quotas and rate limits, and Stripe billing with webhook signature verification.

EDUCATION

BS Information Technology | University of Santo Tomas, Manila, Philippines Aug 2021 - Jun 2025

- **Summa Cum Laude** with a GWA of 1.169 on a 1.0 scale. **Rank 1 in the IT Department**, Batch 2025, Rank 2 overall in the College of Information and Computing Sciences, and Dean's Lister every term from 2021 to 2025.
- Top 3 Best Capstone Project for a full stack web based graduate career tracer, taken from requirements gathering through deployment.

CERTIFICATIONS

Anthropic Claude with the Anthropic API (Jun 2026), Claude Code in Action (May 2026), Introduction to Model Context Protocol (May 2026), Introduction to Agent Skills (May 2026), Claude 101 (Feb 2026), AI Fluency: Framework and Foundations (Jan 2026)

PMI Generative AI Overview for Project Managers (Feb 2026), Free Introduction: PMI Certified Professional in Managing AI, PMI-CPMAI (Feb 2026)

Cisco CCNAv7 Enterprise Networking, Security and Automation (Jan 2024), Switching, Routing and Wireless Essentials (Aug 2023), Introduction to Networks (Jan 2023)

TECHNICAL SKILLS

AI and Agentic Anthropic Claude API, Claude Agent SDK, Claude Code, Model Context Protocol (MCP), sub agents, skills and hooks, OpenAI API, Google Gemini, Vertex AI, Dialogflow CX, Microsoft Copilot, Copilot Studio, LangGraph, RAG and retrieval pipelines, prompt engineering, tool and function calling, structured and JSON schema output, evals and golden task regression, groundedness verification, LLM cost and latency budgeting, n8n, Dify

Languages TypeScript, JavaScript, Python, SQL, PHP, Java, Kotlin, C#, C++

Frontend Next.js (App Router, RSC), React 19, Astro, TailwindCSS v4, shadcn/ui, Material UI, Figma, WCAG 2.1 and 2.2 AA accessibility

Backend and Data Node.js, NestJS, Hono, REST and OpenAPI 3.1, tRPC, GraphQL, Zod, Valibot, Drizzle, Convex, PostgreSQL, Neon, Supabase, Cloudflare D1, Firestore, PayloadCMS

Automation and Integration n8n, ConnectWise Automate (RMM), ConnectWise Manage, Neo and Thread (PSA), triage agents, connectors, webhook and event pipelines, scheduled jobs

Cloud and DevOps Vercel, Cloudflare (Workers, Pages, R2, Cron), Google Cloud Platform, Docker, Dokploy, GitHub and GitLab CI/CD, Playwright, Vitest, PostHog, GA4, Search Console, Tag Manager