

Cayman Roden

AI Engineer | Python, FastAPI, LLM applications, agents, and evaluation

310-982-0492 | caymanroden@gmail.com | linkedin.com/in/caymanroden | github.com/ChunkyTortoise | chunkytortoise.github.io | Remote, US-based

Proof: github.com/ChunkyTortoise/jorge_real_estate_bots | github.com/ChunkyTortoise/docextract | github.com/ChunkyTortoise/llm-reviewer-path | github.com/ChunkyTortoise/mcp-server-toolkit

Professional Summary

AI application engineer. Shipped a bilingual SMS qualification system for Acuity Real Estate that processed 500+ inbound leads in a three-month paid engagement. Public eval proof: DocExtract at 95.5% weighted field-level accuracy on a 28-case offline replay and a 202-case corpus (151 golden, 51 adversarial). Reviewers start at [llm-reviewer-path](https://github.com/ChunkyTortoise/llm-reviewer-path), then [mcp-server-toolkit](https://github.com/ChunkyTortoise/mcp-server-toolkit) on GitHub.

Experience

AI Engineer (paid engagement) | Acuity Real Estate | Remote, US | January 2026 to March 2026

- Built and shipped a Claude-powered bilingual (EN/ES) SMS qualification system that processed 500+ inbound leads during a three-month client run, with GoHighLevel updates and warm-lead routing.
- Worked with the broker to define qualification and escalation rules, then built three specialized bots for lead intake, buyer qualification, and seller qualification with cross-bot routing and human handoff.
- Added golden and adversarial fixtures, rubric scoring, and evidence extraction for the SMS flows.
- Delivered 1,700+ tests at handoff with a test strategy, coverage report, runbook for new flows, and an audit of 226 existing GoHighLevel workflows.

Independent engineering (not an employer) | Remote, US | Present

Independent public work, not employment and not Acuity's employer. Selected projects below are independent unless labeled paid.

Service Industry and Operations | Los Angeles and Palm Springs, CA | 2010-2024

Customer-facing service, sales, and bookkeeping roles held while building the engineering practice.

Selected Projects

DocExtract | github.com/ChunkyTortoise/docextract

- Built a document-extraction RAG service with two-pass Claude tool use, pgvector retrieval, and structured outputs.
- Reached 95.5% weighted field-level accuracy on a 28-case offline replay; maintained a 202-case corpus (151 golden and 51 adversarial) with deterministic replay and an 80% CI coverage gate.

llm-reviewer-path | github.com/ChunkyTortoise/llm-reviewer-path

- Built a no-API-key review path covering eval gates, hard-action approval boundaries, retrieval failure modes, and delivery-scoping evidence.
- Packaged a reproducible local setup (`uv sync --group dev,uv run pytest`) so reviewers can inspect engineering decisions and failure handling directly.

mcp-server-toolkit | github.com/ChunkyTortoise/mcp-server-toolkit

- Built a Python MCP server framework with JWT resource-server auth, structured tool outputs, multi-tier caching, and OpenTelemetry tracing with per-span cost attribution.
- Maintained 617 collected tests on the GitHub 0.3.0 tree (2026-08-26 collect-only), a 30-case adversarial injection corpus, and a Python 3.10 to 3.12 CI matrix.
- GitHub is the supported artifact. PyPI 0.1.0 is a different older package, not this tree.

Technical Skills

Python, FastAPI, PostgreSQL, Redis, SQLAlchemy, Pydantic, pytest, Docker, GitHub Actions, Claude API, OpenAI API, Gemini, RAG, embeddings, pgvector, MCP, tool calling, structured outputs, golden datasets, adversarial evaluation, OpenTelemetry, human handoff