Oren Segal

Applied AI Engineer · Agents, Evals and Full-Stack

Brooklyn, NY | 917-232-4270 | orenssegal@gmail.com | U.S. Permanent Resident, no sponsorship needed
linkedin.com/in/oren-segal | github.com/OrenSegal | orensegal.github.io

PROFESSIONAL SUMMARY

Applied AI engineer who builds LLM products and measures them with evals and cost controls. Founder of Shelfie, a vision-powered iOS kitchen app. Also publishes open-source agent and eval tooling. Previously five years as a Senior Data Analyst at Reshet 13, building the forecasting and ad-pricing models behind its programming and sales decisions.

WORK EXPERIENCE

Shelfie 2025-present
Founder & Full-Stack AI Engineer · vision-powered kitchen platform for iOS, pre-launch New York

- Built the Swift app and Postgres backend behind the kitchen tour, which turns a camera walk-through into a live inventory, adapts recognition to each household's corrections and picks dinner from what is on the shelf and expiring soon.
- Cropped items on device and sent them to the cloud model in batches of up to nine, so per-item input cost falls up to ~9x by design.
- Capped AI spend with a daily gateway budget that grows with paid subscribers above a fixed floor and sheds free traffic first, and gated scan changes on replay evals.

Cohort NYC 2026-2026
Growth Engineer (volunteer) · NYC community nonprofit in formation New York

- Optimized the event funnel from visibility to attendance with an agent-driven system for copy, distribution targeting and core audiences; over two months followers grew 61% and attendance 55% (before and after, no control group).

Career break · U.S. relocation and full-time AI/ML retraining 2023-2025

Reshet 13 2018-2023
Senior Data Analyst · national TV network Tel Aviv

- Saved about 70 work hours a week by moving Excel reporting to pandas and SQL pipelines, adopted across the network and its news company, and mentored four junior analysts through the rollout.
- Replaced legacy pricing heuristics with a regression model that found underpriced primetime ad slots, and took it live with the commercial team in FY2022 (gross ad revenue rose 17% that year; no holdout isolated the model's share).
- Built the audience analytics system in Power BI, replacing gut-based calls with cleaned viewership data, and the forecasting model that became the standard input for content acquisition and greenlight decisions.

PROJECTS

Open-source agent and eval tools

- **rootcause:** incident-response agent with an approval gate, graded on 25 incidents with known causes. Across 525 runs from Haiku to Opus, cost per solved incident varied 25x while pass rates of 77 to 84% showed no detectable difference at that sample size; in a separate 300-run test, adding worker agents tripled cost with no detected accuracy gain.
- **litmus:** mutation testing for agent evals that shows which prompt changes they miss and flags graders and LLM-judge rubrics that cannot fail. 2 of 6 eval cases across three plugins still passed with the plugin removed.
- **sous:** Claude Code harness that blocks destructive commands and secret reads and fails test runs that ran nothing.

SKILLS

AI engineering: LLM agents and tool calling, multi-agent orchestration, LangGraph, evals and LLM-judge checks, prompt-injection testing, RAG, MCP, Claude Code plugins, hooks and guardrails, OpenTelemetry tracing
Software: Python, TypeScript, React, Next.js, Swift, SQL on Postgres and pgvector, CI and testing, Docker
Data and ML: regression and forecasting, bootstrap intervals and power analysis, data pipelines, Power BI, pandas, scikit-learn, TensorFlow

EDUCATION & CERTIFICATIONS

B.Sc. Information Systems, The Academic College of Tel Aviv-Yaffo 2014-2018
Certifications: Duke MLOps & LLMOps 2025 · Agent Engineering Bootcamp 2025 · DeepLearning.AI Deep Learning 2024