

O. FELIX AMORUWA

famoruwa@berkeley.edu | 909-731-9011 | felixamoruwa.info

Applied AI architect and technical product leader with 12+ years guiding enterprise customers from discovery to deployment — most recently building and shipping production Claude (MCP SDK) integrations and a multi-agent orchestration framework spanning real estate, financial services, insurance, and health. Designed enterprise LLM evaluation frameworks (aeval: adversarial safety testing, bootstrap CI, automated safety gates) and RL post-training workbenches benchmarking GRPO/DPO across TRL, VeRL, OpenRLHF, and NeMo RL. Scaled developer platform adoption to 675M+ engagements at Intuit, delivered C-suite technical strategy, and translated complex AI architectures for diverse stakeholders as a De Anza adjunct faculty member. NeurIPS published researcher; B.S. Computational Engineering, UC Berkeley; M.B.A., Carnegie Mellon.

PROFESSIONAL EXPERIENCE

STREAMIO AI | Oakland, CA

Founder & CEO — Applied AI Architect (Claude API, Multi-Agent, Cross-Industry) 09/24 – Present

- Architected and shipped production Claude (MCP SDK) integrations across four enterprise verticals — real estate, financial markets, insurance, and health/dental — designing scalable agentic workflows that translate complex business requirements into Claude-powered solutions; built OpenClaw multi-agent orchestration framework (gateway protocol, subagent delegation, profile management, session switching) coordinating AI agent workflows end-to-end.
- Built Vantage AI Job Search & Interview Prep platform (iOS/macOS) with 40+ agent tools over a shared FastAPI backend — implemented STAR behavioral evaluation with an automated AI judge, LeetCode-catalog coding drills, and per-role artifact generation; designed SSE streaming architecture, idempotent endpoints, and async job polling for reliable enterprise-grade delivery.
- Delivered live 1:1 AI coaching sessions (Daily/Tavus video, camera/mic capture) and MAS-safe embedded video calls with native macOS 15 system-picker fallback; shipped App Review-compliant monetization (Apple IAP + Superwall, Kinde OAuth + Apple Sign-In) with explicit AI data-sharing consent gating and Playwright/Jest release gates.
- Led 0-to-1 product strategy, technical architecture, and go-to-market execution — conducted customer discovery interviews across multiple industries, iterated on product-market fit, and established influencer partnerships; built and shipped signed/notarized cross-platform builds (macOS DMG, Linux .deb) with native ScreenCaptureKit via Swift.

FINTELLECT AI | Oakland, CA

Founder & CEO — Enterprise LLM Architect, Financial Services 09/24 – Present

- Architected a production RAG retrieval pipeline (ChromaDB vector store) with multi-provider LLM orchestration (Claude, GPT-4, Gemini), fallback routing, structured-output validation, and token-budget optimization — deployed 13 specialized domain AI advisors delivering guided, context-aware financial advisory at scale.
- Shipped a regulated financial services AI platform (iOS/Android/Web) with paper trading on a \$100K virtual portfolio (market/limit/stop/stop-limit orders, P&L, AI portfolio analysis) on live Alpaca market data; navigated App Review compliance across desktop and mobile including Apple IAP, App Store Server Notifications, and consent-gated data flows.
- Designed and delivered enterprise-grade security architecture: MAS sandbox entitlements, no remote debug ports, strict IPC allowlists, Kinde OAuth + Sign in with Apple with custom deep-link protocols — directly applicable to Anthropic's Project Glasswing cross-industry security initiative.
- Led customer discovery on trader feedback, established influencer partnerships, and executed go-to-market through App Store launch; translated complex LLM architecture decisions into business-value narratives for non-technical stakeholders.

INTUIT | Mountain View, CA

Staff Product Manager — Developer Frameworks & Platform Infrastructure 06/21 – 09/24

- Scaled ICE platform to 675M+ engagements in FY23 (275% YoY growth) across QuickBooks, TurboTax, Mint, Mailchimp, and Credit Karma; drove throughput from 6K to 50K TPS via rSocket migration supporting ~1.5M concurrent connections at sub-25ms TP99 — enterprise platform scale directly relevant to Claude's accelerating adoption.
- Delivered ICE Self-Service platform (DevPortal, GitOps config, ICE Playground), reducing developer onboarding from 2–3 weeks to minutes in pre-prod and <24 hours for production, while mitigating \$1M+ in projected opex growth — a model for accelerating enterprise Claude adoption journeys.
- Extended Java and Python SDK Starter Kits with scaffolding templates, build configurations (Gradle/Maven), testing frameworks, and CI/CD integration — empowering developers to go from zero to production-ready microservice in minutes; led Mailchimp GCP-to-AWS migration delivering Golang template and updated DevPortal documentation.
- Conducted enterprise-wide Service Language Assessment across 9 languages, analyzing usage data and developer feedback to inform strategic investment decisions presented to CTO; worked with telemetry and SQL/BigQuery to prioritize developer pain points across ~20 mobile apps and 30+ product SKUs.
- Implemented ICE Presence in async chat generating \$480K/month in additional invoicing; deployed Background-to-Foreground Messaging on iOS/Android with <100ms latency. DeveloperWeek 2022 speaker.

SPLUNK INC. | San Francisco, CA

Senior Product Manager — Search Orchestration 01/19 – 01/21

- Served as primary technical advisor to enterprise beta customer (Assurance), building a mirrored Enterprise topology for benchmark testing and achieving up to 10x query performance improvements in Splunk Cloud Services — translating complex technical findings into actionable architecture recommendations.
- Owned Search Service (Go microservices), Search Catalog (PostgreSQL), and Splunk Processing Language (SPL/SPL2); delivered Scheduler Service end-to-end in ~4 months, enabling scheduled search for first-party applications; demoed at Splunk .conf19.
- Designed repeatable RICE-based prioritization framework balancing internal partner, third-party developer, and Fortune 500 customer requirements across 3 microservice backlogs. Splunk .conf18 Speaker.

KAISER PERMANENTE | Pleasanton, CA

SOA Technical Product Manager 08/12 – 12/18

- Led enterprise rollout of Splunk Logging-as-a-Service (1.7 TB daily, 200+ internal enterprise customers) and ITSI Application Monitoring-as-a-Service in a regulated healthcare environment; built Redis/XC10 caching for scalability and fault tolerance across the enterprise.
- Led demand forecasting and capacity planning for SOA product suite across multiple datacenters, translating technical capacity constraints into business-aligned investment roadmaps for senior stakeholders.

IBM | Pittsburgh, PA / Foster City, CA / RTP, NC

Software Engineer — Business Intelligence Products 07/05 – 03/10

- Resolved multiple \$1–5M high-severity customer escalations for enterprise BI software; led cross-functional root cause analysis and customer-facing resolution, improving case resolution time by 20% — foundational enterprise customer-facing technical advisory experience.

AI/ML RESEARCH & PROJECTS

aeval — AI Model Evaluation Platform (2025–2026)

- Built local-first LLM evaluation platform with 5 core eval types (factuality, reasoning, instruction-following, safety, code generation), adversarial safety testing with refusal detection, and data contamination detection via SHA-256 hashing — directly applicable to helping enterprise customers develop evaluation frameworks for Claude.
- Statistical rigor: bootstrap confidence intervals, Welch's t-test, Cohen's d effect size, saturation detection; CI/CD integration with regression detection and automated safety gates. Stack: FastAPI orchestrator, TimescaleDB, Redis job queue, Next.js dashboard, Ollama.

RL Workbench — Post-Training RL Platform (2026)

- Built 3-phase RL post-training workbench covering the full RLHF/DPO pipeline: Reward Lab for designing and A/B testing reward functions (RLVR, learned, hybrid) across 4 datasets (GSM8K, MATH, HumanEval, UltraFeedback), Playground for real TRL-powered GRPO/DPO training with live SSE metric streaming on Apple Silicon (MPS) or CUDA, and Arena for head-to-head framework benchmarking (TRL, VeRL, OpenRLHF, NeMo RL) with GPU passthrough in Docker containers.
- Implemented 12 RL algorithms (PPO, GRPO, DAPO, REINFORCE, REINFORCE++, RLOO, DPO, SimPO, IPO, KTO, ORPO, SPPO) with algorithm-specific metric profiles, cross-tab workflow lineage tracking, and standardized throughput/memory/convergence benchmarking across frameworks.

BRAIN — Protein Structure Prediction ML Platform (UC Berkeley 2004; rewritten 2026)

- Production ML platform in PyTorch with 5 neural architectures (feedforward, GRU, Transformer, ESM-2, multi-task), MLflow experiment tracking, Optuna HPO, and FastAPI serving — 823 automated tests, Docker orchestration (6 containers); original 2004 system hand-coded in C++ with custom BPTT, rewritten to span 413 to 8B parameters (19M-fold scale increase).

NeurIPS 2014 — Montreal, Canada

- Accepted paper on artificial neural networks for protein secondary structure prediction; computational analysis of RNA/DNA sequences under Dr. Steven Holbrook at Lawrence Berkeley National Laboratory.

EDUCATION

Carnegie Mellon University, Silicon Valley

Master of Science — Software Management Mountain View, CA (08/16)

Carnegie Mellon, Tepper School of Business

Master of Business Administration — Finance, Quantitative Analysis, Economics Pittsburgh, PA (03/12)

UC Berkeley, College of Engineering & Haas School of Business

B.S. — Computational Engineering Science · Certificate — Corporate Finance for Financial Engineering Berkeley, CA (12/05)

TEACHING EXPERIENCE

DE ANZA COLLEGE | Cupertino, CA

Adjunct Faculty — Computer Information Systems Department 04/18 – Present

- Teach cloud computing (AWS, GCP), data analytics, Java programming, ethical hacking, digital forensics, and PL/SQL to diverse student audiences — translating complex technical concepts into accessible instruction across skill levels, directly reflecting the JD's emphasis on teaching, mentoring, and communicating to diverse stakeholders.
- Courses: CIS 35A/B, 36A/B — Java Programming; CIS 64C/64G — PL/SQL; CIS 44F — Data Analytics; CIS 64E — Large Scale Cloud Computing (AWS, GCP); CIS 102 — Ethical Hacking; CIS 104 — Digital Forensics; CIS 95F — Managing Cloud Projects.

ADDITIONAL INFORMATION

- FAA Commercial Drone License; FAA Private Pilot License (intended CFI).
- O'Reilly & OER Publications: Data Analytics and Managing Cloud Projects (forthcoming Summer 2026).
- First Tee Assistant Golf Coach — San Francisco Chapter (2012–Present), Pittsburgh Chapter (2011–2012).
- Interests: Classical & Jazz Saxophone (Alto/Tenor/Baritone), Triathlon, Golf, Intramural Basketball & Soccer.