

O. FELIX AMORUWA

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

Technical Product Manager with 12+ years building developer-facing platforms, SDKs, and AI inference infrastructure at scale — scaling ICE to 675M+ engagements and 50K TPS at Intuit, shipping developer SDKs and self-service DevPortals that cut onboarding from weeks to minutes, and architecting multi-agent orchestration frameworks and RL inference workbenches from 0 to 1. Deep hands-on depth in agent frameworks, container-based AI inference pipelines, and open-source developer tooling. NeurIPS published researcher; UC Berkeley BS Computational Engineering.

PROFESSIONAL EXPERIENCE

INTUIT | Mountain View, CA

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

- Owned developer experience roadmap for SDK Starter Kits and ICE self-service platform — extended Java and Python SDKs with scaffolding templates, build configs (Gradle/Maven), testing frameworks, and CI/CD integration, enabling developers to go from zero to production-ready microservice in minutes.
- Achieved 275% YoY growth in ICE engagements, scaling to 675M+ in FY23 across QuickBooks, TurboTax, Mint, Mailchimp, and Credit Karma; scaled throughput from 6K to 50K TPS via rSocket migration supporting ~1.5M concurrent connections with sub-25ms TP99.
- 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.
- Conducted enterprise-wide Service Language Assessment across 9 languages, analyzing usage data and developer feedback to inform strategic platform investment decisions presented to CTO — translating developer pain into product strategy at executive level.
- Built Asterias, a declarative asset lifecycle management platform with GraphQL API; worked closely with telemetry and usage data (SQL, BigQuery) to prioritize developer pain points across ~20 mobile apps and 30+ product SKUs.
- Led Mailchimp GCP-to-AWS migration for MSaaS, delivering Golang service template, MySQL persistence integration, and updated DevPortal documentation; initiated MSaaS Drift Detection program with Java JAR library scanning Git repos for configuration drift.
- 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.

STREAMIO AI | Oakland, CA

Founder & CEO — Developer Platform & Agentic Workflows (streamio.ai) 09/24 – Present

- Architected OpenClaw multi-agent orchestration framework with gateway protocol, subagent delegation, profile management, and session switching — enabling coordinated agentic workflow support across real estate, insurance, health/dental, and financial markets verticals.
- Shipped MCP SDK integration exposing screen capture tools to AI coding assistants and external agent frameworks; built embedded terminal (xterm.js/node-pty, 20+ slash commands) as a developer-facing CLI surface within the desktop application.
- Built production Electron + React + TypeScript desktop application with 100+ components, Redux Toolkit state management, and native macOS/Linux support — delivering real-time AI inference of screen captures using Claude, with contextual conversation and automated report generation.
- Engineered real-time HLS livestreaming pipeline: multi-stream canvas compositing (up to 9 concurrent sources at 30fps), FFmpeg transcoding, WebSocket layer, and local HTTP playback server; built cross-platform iOS app (React Native/Expo) with on-device FFmpeg transcoding and Kinde OAuth 2.0.
- Led 0-to-1 product strategy, AI development, and go-to-market execution including customer discovery interviews and iterative product refinement.

FINTELLECT AI | Oakland, CA

Founder & CEO — AI Inference & LLM Orchestration Platform (fintellect.ai) 09/24 – Present

- Architected multi-provider LLM inference orchestration pipeline (Claude, GPT-4, Gemini) with RAG retrieval (ChromaDB), fallback routing, structured output validation, and token budget optimization — delivering AI inference at scale for retail investor workflows.
- Built domain-specific conversational AI agents embedded in mobile and web app, each scoped to distinct financial focal points, delivering guided context-aware advisory interactions; executed go-to-market through App Store launch.

SPLUNK INC. | San Francisco, CA

Senior Product Manager — Search Orchestration & Cloud Services 01/19 – 01/21

- Owned product roadmap for Search Service (Go microservices), Search Catalog (PostgreSQL metadata service), and Splunk Processing Language (SPL/SPL2) for Splunk Cloud Services — built PRDs, user stories, and acceptance criteria for a cloud-native developer platform.
- Delivered Scheduler Service end-to-end in ~4 months, enabling scheduled search capabilities for first-party applications; demoed at Splunk .conf19; led query performance initiative achieving up to 10x improvements for Fortune 500 design-partner (Assurance).
- Designed repeatable RICE-based prioritization framework for 3 microservice backlogs, balancing internal partner, third-party developer, and Fortune 500 customer requirements. Splunk .conf18 Speaker.

KAISER PERMANENTE | Pleasanton, CA

SOA Technical Product Manager 08/12 – 12/18

- Led development and enterprise rollout of Splunk Logging-as-a-Service (1.7 TB daily volume, 200+ internal enterprise customers) and ITSI Application Monitoring-as-a-Service.
- Built Redis and XC10 caching capability across the enterprise to address scalability, fault tolerance, and data redundancy; led demand forecasting and capacity planning for SOA product suite across multiple datacenters.

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% across the Support Organization.

AI/ML RESEARCH & PROJECTS

RL Workbench — Post-Training RL & AI Inference Benchmark 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 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, DPO, SimPO, IPO, KTO, ORPO, SPPO, REINFORCE++, RLOO) with standardized throughput/memory/convergence benchmarking across inference frameworks — directly analogous to NVCF inference workload benchmarking.

aeval — AI Model Evaluation Platform (2025–2026)

- Built local-first evaluation platform with 5 core eval types (factuality, reasoning, instruction-following, safety, code generation), adversarial safety testing, and CI/CD integration with regression detection and automated safety gates. Stack: FastAPI orchestrator, TimescaleDB, Redis job queue, Next.js dashboard, Ollama.

AutoEval — Automated Visual Evaluation for Robot Model Training (2025)

- Built zero-integration automated visual evaluation system for robot model training using multimodal AI inference (Claude/GPT-4V) for spatial reasoning on prediction frames — reducing evaluation cycles from 72 hours to ~4 minutes with structured PASS/FAIL reports and confidence scoring.

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

- NeurIPS 2014 accepted paper; production ML platform in PyTorch with 5 neural architectures (feedforward, GRU, Transformer, ESM-2, multi-task), MLflow experiment tracking, Optuna HPO, FastAPI serving — 823 automated tests, Docker orchestration (6 containers), 413 to 8B parameters (19M-fold scale).

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 — Concentrations: Finance, Quantitative Analysis, Economics Pittsburgh, PA (03/12)

UC BERKELEY, COLLEGE OF ENGINEERING & HAAS SCHOOL OF BUSINESS

Bachelor of Science — Computational Engineering Science Berkeley, CA (12/05)

Certificate — Corporate Finance for Financial Engineering

TEACHING EXPERIENCE

DE ANZA COLLEGE | Cupertino, CA

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

- CIS 35A/B, 36A/B — Java Programming; CIS 64C/G — 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