

# O. FELIX AMORUWA

famoruwa@berkeley.edu | 909-731-9011 | [felixamoruwa.info](mailto:felixamoruwa.info)

Technical Product Manager with 12+ years building developer-facing APIs, SDKs, and ML platform infrastructure — scaling ICE to 675M+ engagements and 50K TPS at sub-25ms TP99 (Intuit), and shipping a full RL post-training workbench benchmarking GRPO/DPO across TRL, VeRL, OpenRLHF, and NeMo RL. Deep builder credibility in production inference pipelines, multi-provider LLM orchestration, and multi-component AI workflows. NeurIPS published researcher; B.S. Computational Engineering Science, UC Berkeley.

## PROFESSIONAL EXPERIENCE

---

**INTUIT** | Mountain View, CA

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

- Scaled ICE platform to 675M+ engagements in FY23 across QuickBooks, TurboTax, Mint, Mailchimp, and Credit Karma — 275% YoY growth — by driving throughput from 6K to 50K TPS via rSocket migration supporting ~1.5M concurrent connections at sub-25ms TP99.
- Delivered ICE Self-Service platform (DevPortal, GitOps config, ICE Playground) that reduced developer onboarding from 2–3 weeks to minutes in pre-prod and <24 hours for production, while mitigating \$1M+ in projected opex growth.
- 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.
- 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.
- Conducted enterprise-wide Service Language Assessment across 9 languages, analyzing usage data and developer feedback to inform strategic language investment decisions presented to CTO.
- 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 — streamio.ai (macOS, Linux, iOS & Web)* 09/24 – Present

- Implemented OpenClaw multi-agent orchestration framework with gateway protocol and subagent delegation — enabling coordinated multi-component AI workflows analogous to Baseten's Chains initiative across real estate, insurance, and financial markets.
- Engineered real-time HLS inference pipeline: multi-stream canvas compositing (up to 9 concurrent sources at 30fps), FFmpeg transcoding, WebSocket communication layer, and local HTTP playback server — production-grade async inference architecture.
- Built production Electron + React + TypeScript desktop application with 100+ components, Redux Toolkit state management, and native macOS/Linux support — delivering real-time AI analysis via Claude (MCP SDK) with contextual conversation workflows.
- Shipped cross-platform builds (macOS DMG with code signing/notarization, Linux .deb) with embedded terminal (xterm.js/node-pty, 20+ slash commands) and MCP server exposing screen capture tools to AI coding assistants.
- 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 — fintellectai.io (iOS/Android/Web)* 09/24 – Present

- Architected multi-provider LLM orchestration (Claude, GPT-4, Gemini) with fallback routing, structured output validation, and token budget optimization — hands-on production model serving across frontier models.
- Built RAG retrieval pipeline with ChromaDB vector store and domain-specific conversational AI agents scoped to distinct financial focal points, delivering context-aware advisory interactions at scale.
- Executed go-to-market through App Store launch; led customer discovery efforts and established partnerships with industry influencers to drive product-led growth.

**SPLUNK INC.** | San Francisco, CA

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

- Owned Search Service (Go microservices), Search Catalog (PostgreSQL metadata service), and Splunk Processing Language (SPL/SPL2) — built roadmaps, PRDs, and acceptance criteria for Splunk Cloud Services.
- Delivered Scheduler Service end-to-end in ~4 months, enabling asynchronous scheduled search capabilities for first-party applications; demoed at Splunk .conf19.
- Led query performance optimization for beta customer (Assurance), achieving up to 10x improvements in SCS search — building mirrored Enterprise topology for benchmark testing. 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 ITS! Application Monitoring-as-a-Service.
- Built Redis caching capability across the enterprise to address scalability, fault tolerance, and data redundancy; led demand forecasting and capacity planning 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 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, 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 — directly analogous to Baseten Loops.

- 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.

#### **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 — full observability and management experience for model inference.
- Stack: FastAPI orchestrator, TimescaleDB, Redis job queue, Next.js dashboard, Ollama; statistical rigor via bootstrap confidence intervals, Welch's t-test, and Cohen's d effect size.

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

- Built 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); spans 413 to 8B parameters.
- NeurIPS 2014 accepted paper on artificial neural networks for protein secondary structure prediction.

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

- Built automated visual evaluation system repurposing screen capture + HLS streaming + multimodal AI pipeline to score model outputs against natural-language rubrics — reducing evaluation cycles from 72 hours to ~4 minutes; zero-integration architecture captures from any visualization tool (RViz, Matplotlib).

## **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 — Introduction to PL/SQL; CIS 44F — Introduction to Data Analytics; CIS 64E — Fundamentals of 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