

O. FELIX AMORUWA

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

Principal-level AI product leader with 15+ years translating applied research into enterprise-scale products — from a NeurIPS-published neural network (2004) to RL post-training workbenches benchmarking GRPO/DPO across TRL, VeRL, OpenRLHF, and NeMo RL today. Founded two AI-native companies, shipped agentic architectures with multi-agent orchestration and LLM/generative model pipelines, and scaled developer infrastructure to 675M+ engagements at Intuit. Deep technical fluency across LLMs, model alignment, physical AI evaluation, and distributed inference — with a track record engaging enterprise customers, cloud partners, and research teams from zero-to-one through commercialization.

PROFESSIONAL EXPERIENCE

STREAMIO AI | Oakland, CA

Founder & CEO — Agentic AI Platform (macOS, Linux, iOS & Web) 09/24 – Present

- Architected and shipped OpenClaw multi-agent orchestration framework — gateway protocol, subagent delegation, profile management, and session switching — enabling coordinated agentic AI workflows across real estate, insurance, health/dental, and financial markets; directly demonstrates agentic software development at production scale.
- Built production Electron + React + TypeScript desktop application with 100+ components and native macOS/Linux support — providing real-time multimodal AI analysis (Claude MCP SDK) with contextual reasoning, Pine Script generation, and automated CMA reports via Redfin/Zillow APIs.
- Engineered real-time HLS livestreaming pipeline: multi-stream canvas compositing (9 concurrent sources at 30fps), FFmpeg transcoding, WebSocket communication layer — enabling live broadcast from desktop and mobile with sub-second latency.
- Shipped cross-platform iOS app (React Native/Expo) with dual-source capture, on-device FFmpeg transcoding, Picture-in-Picture compositing, and Kinde OAuth 2.0 — demonstrating full-stack zero-to-one product execution from research through commercialization.
- Led 0-to-1 product strategy, applied AI development, and go-to-market execution including customer discovery and iterative product refinement with AI-native customers.

FINTELLECT AI | Oakland, CA

Founder & CEO — AI-Native Financial Platform (iOS/Android/Web) 09/24 – Present

- Architected RAG retrieval pipeline with ChromaDB vector store, multi-provider LLM orchestration (Claude, GPT-4, Gemini) with fallback routing, structured output validation, and token budget optimization — delivering enterprise-grade reliability for AI-native retail investors.
- Built domain-specific conversational AI agents scoped to distinct financial focal points, delivering guided, context-aware advisory interactions at scale; executed go-to-market through App Store launch and strategic influencer partnerships.

INTUIT | Mountain View, CA

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

- Scaled ICE developer platform to 675M+ engagements in FY23 across QuickBooks, TurboTax, Mint, Mailchimp, and Credit Karma; drove 275% YoY growth and scaled throughput from 6K to 50K TPS via rSocket migration supporting ~1.5M concurrent connections with sub-25ms TP99.
- 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; a developer infrastructure product adopted at enterprise scale.
- 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 (Java, Python, Kotlin, Go, TypeScript, Scala, PHP, C++, Groovy), analyzing usage data and developer feedback to inform strategic investment decisions presented to CTO.
- Led Mailchimp GCP-to-AWS cloud migration for MSaaS; initiated Drift Detection program with Java JAR library scanning Git repos for configuration drift and OpenRewrite-based remediation roadmap.
- 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 & Cloud Services 01/19 – 01/21

- Owned Search Service (Go microservices), Search Catalog (PostgreSQL), and Splunk Processing Language (SPL/SPL2) — defining roadmaps and PRDs for Splunk Cloud Services (SCS) serving Fortune 500 enterprise customers.
- Delivered Scheduler Service end-to-end in ~4 months, enabling scheduled search for first-party applications; led query performance optimization achieving up to 10x improvements for beta enterprise customer. Splunk .conf18/.conf19 speaker.
- Designed RICE-based prioritization framework across 3 microservice backlogs, balancing internal partner, third-party developer, and Fortune 500 customer requirements.

KAISER PERMANENTE | Pleasanton, CA

SOA Technical Product Manager — Security & Observability Infrastructure 08/12 – 12/18

- Led enterprise rollout of Splunk Logging-as-a-Service (1.7 TB daily volume, 200+ internal enterprise customers) and ITSI Application Monitoring-as-a-Service — delivering security observability infrastructure at scale.
- Built enterprise caching capability using Redis and XC10 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 improving case resolution time by 20% across the Support Organization.

AI/ML RESEARCH & PROJECTS

RL Workbench — Post-Training RL & Model Alignment Platform (2026)

- Built 3-phase RL post-training workbench covering the full RLHF/alignment pipeline: Reward Lab for designing and A/B testing reward functions (RLVR, learned, hybrid) across GSM8K, MATH, HumanEval, and 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 across TRL, VeRL, OpenRLHF, and 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 — directly advancing state-of-the-art model alignment research.

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

- Built automated visual evaluation system for physical AI robot model training — repurposing screen capture + HLS streaming + multimodal AI pipeline to score model outputs (grasp poses, segmentation maps, bounding boxes) against natural-language rubrics, reducing evaluation cycles from 72 hours to ~4 minutes.
- Zero-integration architecture captures from any visualization tool (RViz, Matplotlib, custom dashboards); multimodal AI (Claude/GPT-4V) performs spatial reasoning on prediction frames, generating structured PASS/FAIL reports with confidence scores and edge-case flagging.

aeval — AI Safety & Model Alignment Evaluation Platform (2025–2026)

- Built local-first evaluation platform with 5 core eval types (factuality, reasoning, instruction-following, safety, code generation), adversarial safety testing with refusal detection, and data contamination detection — directly addressing privacy-preserving ML and model alignment research priorities.
- 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.

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

- NeurIPS 2014 accepted paper on artificial neural networks for protein secondary structure prediction — published applied research translated into production ML platform.
- 2026 rewrite: PyTorch platform with 5 neural architectures (feedforward, GRU, Transformer, ESM-2, multi-task), MLflow experiment tracking, Optuna HPO, FastAPI serving, 823 automated tests, Docker orchestration — scaling from original 413-parameter C++ BPTT system to 8B parameters (19M-fold increase).

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

- Courses: Java Programming (CIS 35A/B, 36A/B); Introduction to PL/SQL (CIS 64C/G); Introduction to Data Analytics (CIS 44F); Fundamentals of Large Scale Cloud Computing — AWS, GCP (CIS 64E); Ethical Hacking (CIS 102); Digital Forensics and Hacker Investigation (CIS 104); Managing Cloud Projects (CIS 95F).

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