

O. FELIX AMORUWA

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

Staff-level Technical PM with 12+ years shipping ML/AI infrastructure and developer platform products at enterprise scale — from scaling a data platform to 675M+ engagements and 50K TPS (Intuit) to building RL post-training workbenches that benchmark GRPO/DPO across TRL, VeRL, OpenRLHF, and NeMo RL today. Deep hands-on ML background: NeurIPS-published researcher, PyTorch/MLflow practitioner, and builder of production model training, evaluation, and RAG/vector search pipelines. Former software engineer (IBM, UC Berkeley) with a track record of translating enterprise ML team pain points into platform capabilities that accelerate the path from experimentation to production.

PROFESSIONAL EXPERIENCE

INTUIT | Mountain View, CA

Staff Product Manager — AI/ML Platform Infrastructure & Developer Frameworks 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 data infrastructure at Databricks scale.
- Delivered ICE Self-Service platform (DevPortal, GitOps config, ICE Playground), compressing developer onboarding from 2–3 weeks to minutes in pre-prod and <24 hours for production, while mitigating \$1M+ in projected opex growth — directly reducing adoption bottlenecks for platform consumers.
- Extended Java and Python SDK Starter Kits with scaffolding templates, build configurations (Gradle/Maven), testing frameworks, and CI/CD integration — enabling developers to go from zero to production-ready microservice in minutes across 20+ mobile apps and 30+ product SKUs.
- Worked closely with telemetry and usage data (SQL, BigQuery) to prioritize developer pain points; built Asterias, a declarative asset lifecycle management platform with GraphQL API, informing strategic investment decisions presented to CTO.
- Led Mailchimp GCP-to-AWS migration for MSaaS, delivering Golang service template, MySQL persistence integration, and updated DevPortal documentation; initiated MSaaS Drift Detection program using Java JAR library to scan Git repos and OpenRewrite for remediation.
- 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 & Data Infrastructure 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, directly analogous to owning ML pipeline and query infrastructure at scale.
- Led query performance optimization initiative for enterprise beta customer (Assurance), building mirrored topology for benchmark testing and achieving up to 10x performance improvements in SCS search — the kind of deep infrastructure work Databricks' AI Platform demands.
- Delivered Scheduler Service end-to-end in ~4 months (June–Oct 2019), enabling scheduled search capabilities for first-party applications; demoed at Splunk .conf19.
- Designed repeatable RICE-based prioritization framework across 3 microservice backlogs, balancing Fortune 500 customer requirements with internal partner and third-party developer needs. Splunk .conf18 Speaker.

STREAMIO AI | Oakland, CA

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

- Built OpenClaw multi-agent orchestration framework with gateway protocol, subagent delegation, and session management — enabling coordinated LLM agent workflows across multiple industry verticals; directly maps to Databricks Agent infrastructure and LLM serving architecture.
- Engineered real-time AI analysis pipeline using Claude (MCP SDK) with screen capture, contextual conversation, and automated report generation — production LLM infrastructure built and operated end-to-end.
- Shipped cross-platform desktop and mobile builds with native OS integrations, embedded terminal (20+ slash commands), and MCP server exposing AI tools to coding assistants; led 0-to-1 product strategy, customer discovery, and go-to-market execution.

FINTELLECT AI | Oakland, CA

Founder & CEO — AI 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 — hands-on experience with the vector search and LLM infrastructure Databricks' AI Platform powers at enterprise scale.
- Built domain-specific conversational AI agents and AI-powered macroeconomic analysis tools; executed go-to-market through App Store launch with customer discovery and influencer partnerships.

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 caching layer for scalability and fault tolerance across the enterprise.

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

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

- Resolved multiple \$1–5M high-severity escalations for enterprise BI software; led cross-functional root cause analysis and customer-facing resolution, improving case resolution time by 20% — former SWE foundation enabling credible engagement with world-class ML engineers.

BANK OF AMERICA MERRILL LYNCH | Charlotte, NC

Tech MBA Summer Associate — Global Corporate & Investment Banking T&O 06/11 – 08/11

- Developed Monte Carlo simulation model to optimize DMAIC phase distribution across a \$494M portfolio in Global Technology & Operations.

AI/ML RESEARCH & PROJECTS

RL Workbench — Post-Training RL Platform (2026)

- Built 3-phase RL post-training workbench covering the full RLHF/DPO ML lifecycle: 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 — directly maps to Databricks Mosaic AI model training infrastructure.
- 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 model evaluation platform with 5 core eval types (factuality, reasoning, instruction-following, safety, code generation), adversarial safety testing with refusal detection, and data contamination detection — maps to Databricks model monitoring and production ML systems governance.
- 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)

- Built production ML platform in PyTorch with 5 neural architectures (feedforward, GRU, Transformer, ESM-2, multi-task), MLflow experiment tracking, Optuna HPO, and FastAPI model serving — 823 automated tests, Docker orchestration (6 containers); scales from 413 to 8B parameters.
- NeurIPS 2014 accepted paper on neural networks for protein secondary structure prediction; original 2004 system hand-coded in C++ with custom BPTT.

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

- Built automated visual evaluation system using multimodal AI (Claude/GPT-4V) to score model outputs (grasp poses, segmentation maps, bounding boxes) against natural-language rubrics — reducing evaluation cycles from 72 hours to ~4 minutes via zero-integration screen capture architecture.

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

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