

Dat Nguyen

datng@engineering.upenn.edu | linkedin.com/in/dat-nguyen294 | github.com/whenindan | datnguyen.xyz

EDUCATION

University of Pennsylvania

Master of Science in Computer Science

Philadelphia, PA

Aug 2026 – May 2028

Michigan State University

Bachelor of Science in Computer Science, Minor in Business

East Lansing, MI

Aug 2022 – May 2026

WORK EXPERIENCE

Jack

Software Engineer Intern

New York, NY

May 2026 – Aug 2026

- Automated job discovery and application for **250K+ users** by architecting a distributed pipeline (**Next.js**, **Supabase**, **pg_cron** RPC scheduling) that scrapes **12+ ATS providers** and drives autonomous browser-based applications on candidates' behalf.
- Improved candidate-job match quality by training a **LightGBM** ranking model to predict interview conversion (leakage-safe splits, **PR-AUC**-optimized for a ~8% positive class), shipped as an env-gated re-ranking layer for safe progressive rollout alongside the production matching engine.
- Built an **LLM-driven** resume pipeline (**Azure OpenAI**) with deterministic, hash-based template routing across **5 LaTeX** templates and **LuaLaTeX/XeLaTeX** compilation, using fail-loud validation to guarantee reproducible, ATS-compliant PDFs.

DIJAM

Software Engineer Intern

Remote

May 2025 – Aug 2025

- Enabled seamless subscription management for **1M+ users** and **15K+ DJs** by designing and building a scalable, distributed pricing-tier and payment platform in **Next.js/TypeScript** that handles recurring billing, upgrades, and cancellations.
- Helped **500+ DJs** earn recurring revenue by designing and implementing a **Python/Flask** affiliate system with **PostgreSQL**-backed commission-tracking algorithms and automated payout scheduling.
- Saved **8+ hours/week** of manual bookkeeping by automating invoice and payment reconciliation using the **Gmail API**, applying string-matching algorithms to parse transaction emails and reconcile them against subscription records.

nuoc.solutions

Software Engineering Intern

San Francisco, CA

May 2024 – May 2025

- Delivered real-time water-quality insights to **1,000+ farmers** by designing and building a scalable, map-based monitoring platform (**Next.js**, **FastAPI**, **PostgreSQL**) covering **200+ stations**.
- Enabled **500+ monthly** natural-language irrigation queries by developing an LLM-powered information-retrieval system over historical and seasonal time-series data for non-technical users.
- Scaled the platform to support multi-province citizen reporting by containerizing services with **Docker** and deploying on **Azure** with horizontal autoscaling in a **Unix/Linux** production environment.

HCLTech

Software Engineering Intern

Hanoi, Vietnam

May 2023 – Aug 2023

- Cut manual security-review time by **70%** by designing and building a **GitHub Actions** CI/CD pipeline that runs automated static-analysis (SAST) scans across **10K+ files** on every commit.
- Flagged **150+ high-risk** security events at **95% precision** by developing and training anomaly-detection algorithms (**Isolation Forest**, **One-Class SVM**) on large-scale event-log data.
- Accelerated detection of and response to security incidents by building real-time file-integrity monitoring infrastructure across **200+ production Unix/Linux systems**.

PROJECTS

ACASO | Founder & Lead Engineer | FastAPI, Next.js, PostgreSQL, Cerebras, Groq, Fireworks AI

- Founded and led a team of **4 engineers** building an AI-powered B2B SaaS platform that matches, audits, and reconciles invoices against contract terms for enterprise clients.
- Reduced manual invoice-review time by **70%** and processing costs by **20%** across **5 client companies**.
- Achieved **94%** invoice-matching accuracy by orchestrating multi-provider inference across **Cerebras**, **Groq**, and **Fireworks AI** with confidence-based routing and automated contract validation in **FastAPI/PostgreSQL**.

Mailman | Rust, Python, LanceDB, RAG, Vector Search

- Built a **RAG**-powered email knowledge base that lets users search large mailbox archives using natural language.
- Processed **100K+** emails at **670+** emails/sec by designing a **Rust + Python** ingestion and embedding pipeline.
- Enabled semantic search across entire mailboxes with dual email- and chunk-level embeddings in **LanceDB**.

Real-Time Crypto Trading Platform | Spring WebFlux, Redis Streams, PostgreSQL, Docker

- Built a real-time backend that streams live crypto market data and supports high-volume concurrent trading.
- Handled live market feeds and high concurrency using **Spring WebFlux** and **Redis Streams**.
- Shipped a secure, cloud-ready **REST API** in a fully containerized **Docker** deployment with automated tests.

TECHNICAL SKILLS

Languages: Python, C++, TypeScript, JavaScript, Java, Rust, SQL, Go, Swift, HTML/CSS

Frameworks & Tools: Spring Boot, Spring WebFlux, Next.js, React, Node.js, FastAPI, Flask, Docker, Kubernetes, AWS, Azure, PostgreSQL, Redis, Kafka, Terraform, Git, GitHub Actions, PyTorch, scikit-learn