

Xing Gao

xinggao6@illinois.edu | 302-765-7930 | github.com/happyaaa | linkedin.com/in/xing-gao

EDUCATION

University of Illinois Urbana-Champaign

Master of Computer Science

Champaign, IL

Aug 2024 – Dec 2025

University of Delaware

Bachelor of Science in Computer Science

Newark, DE

Aug 2017 – May 2021

EXPERIENCE

Vicino AI

Software Engineer, Applied AI

Seattle, WA

Jan 2026 – Present

- Built and productionized **multimodal AI generation pipelines** for image, video, and 3D workflows using **Python, FastAPI, Redis, and SQL-backed services**.
- Refactored backend generation architecture by decomposing ambiguous product requirements into reusable **generation pipelines** and service boundaries, improving iteration speed and maintainability across AI workflows.
- Designed **LLM-powered agentic workflows** with prompt optimization, evaluation loops, and selective regeneration to improve output consistency for AI-generated assets.
- Implemented durable background processing with **Redis Streams, async workers, retries, and SSE/polling status updates**, improving reliability and observability for long-running AI jobs in production.
- Iterated on generation workflow behavior and provider/model selection based on user feedback and production usage patterns, improving the quality and usability of AI generation flows.
- Owned production readiness across **API design, monitoring, and AWS CDK/CloudWatch automation** for deployed AI services.

PayPal

Software Engineer @ Beaconfire Solutions

Austin, TX

Oct 2021 – Dec 2023

- Maintained and enhanced large-scale **Java + Spring Boot backend components** for PayPal's global **compliance and sanctions platform**, supporting transaction and account screening, regulatory review workflows, and policy enforcement for a high-stakes payment system.
- Led migration of 5+ core components to **OTEL/SignalFx**, implementing distributed tracing, centralized logging, and fine-grained service metrics to improve observability and speed up on-call debugging.
- Integrated regression and integration test pipelines into **Splunk dashboards**, surfacing real-time service health data for analysts and SREs and improving incident detection across teams.
- Partnered with engineering and compliance stakeholders to deliver backend changes for screening, account review, and operational investigation workflows in a mission-critical regulatory system supporting payment operations across 200+ markets.

ADDITIONAL EXPERIENCE

University of Illinois Urbana-Champaign

Backend Software Engineer – AI Education Platform (RA)

Champaign, IL

Aug 2025 – Dec 2025

- Built a **FastAPI backend** for an AI-assisted platform, exposing **knowledge graph** and analytics APIs to support assignment evaluation and student-level analysis.
- Integrated **LLM-assisted SQL error mapping** and deployed stateless services on **GCP Cloud Run** with JWT auth.

Bili

Software Engineer Intern

Seattle, WA

Jul 2025 – Aug 2025

- Built a **LangGraph**-orchestrated Google Docs/Sheets pipeline with section-level diff/patch, JSON Schema validation, OAuth2 (token refresh, least-privilege), idempotent retries, and audit logs.
- Reduced E2E latency by 30% via async I/O, fan-out/fan-in, API batching, and schema caching; standardized **Pydantic/JSON Schema** contracts with structured logs, trace spans, and adaptive rate limiting.

NaviData AI

Software Engineer Intern (Backend) @ Supstat Inc

New York, NY

May 2025 – Jun 2025

- Developed **FastAPI + Azure SQL + Redis** recommendation engine with **semantic search, collaborative filtering**, and real-time query optimization; reduced P95 latency from 800ms to 200ms.

TECHNICAL SKILLS

Languages: Java, Python, Go, C++, JavaScript, TypeScript

Frameworks: Spring Boot, FastAPI, React, Node.js, gRPC, Hibernate, Express.js

Databases: PostgreSQL, MySQL, MongoDB, Azure SQL, Neo4j, Redis

Cloud/DevOps: AWS, GCP, Azure Cloud, Docker, CloudWatch, Splunk, SignalFx, Git, Postman

Concepts: Applied AI, Multimodal AI, LLM Applications, Agentic Workflows, LLM Evaluation, Async Task Orchestration, Backend System Design, Microservices, Monitoring & Observability, Distributed Tracing