Yun(Zoey) Zeng

+1(347) 397-9896| yzeng018@gmail.com | LinkedIn | GitHub | Brooklyn, NY

EDUCATION

New York University
M.S. in Computer Engineering *Sep 2025 – May 2027*
Relevant Coursework: Data Structures & Algorithms, Operating Systems, Machine Learning, Deep Learning

Beijing Normal University
B.S. in Physics *Sep 2021 – Jun 2025*
Honors: Third Prize, China Undergraduate Mathematical Contest in Modeling; Academic Scholarship

WORK EXPERIENCE

Insta360 **Shenzhen, China**
Software Engineer Intern Jun 2026 – Sep 2026

- Owned end-to-end design and phase-1 launch of a **real-time anti-scalping system (Java, Spring, MySQL, Redis)** for Insta360 global online store**,** reducing scalper orders by **78% i**n week one with **zero false positives**
- Served 6-dimension blocklist checks from Redis at **<50 ms P99 (~1 ms typical)**; synced from MySQL via atomic RENAME rebuilds, with fail-open fallback to protect checkout
- Integrated CAPTCHA verification (Tencent Cloud, Cloudflare Turnstile) into checkout with Redis-based replay protection across Web, H5, and Mini Program
- Rolled out safely via **shadow mode → staged rollout**, measuring false-positive rate before enabling blocking
- Built offline detection pipelines: **DBSCAN clustering over 4M shipping addresses** to surface scalper rings, plus order-interval analysis (coefficient of variation) to flag scripted purchases

Amazing Seasun Games (Kingsoft) **Zhuhai, China**
Software Engineer Intern Dec 2024 – Jan 2025

- Built a **real-time C++ performance monitor for Android** tracking 6 metrics per frame, with sliding-window anomaly detection to auto-flag frame drops
- Added **Jenkins CI regression gates (>10% FPS drop)** that block merges on every build
- Profiled render hotspots with Optick and optimized the rendering pipeline, **cutting render-thread time by 33%**

PROJECTS

Document AI: Full-Stack RAG & MCP-Based Document Q&A Agent Aug 2026 – Present

- Created an end-to-end document agent using **Next.js, React, and TypeScript**, utilizing Tailwind CSS and AG-UI to manage state synchronization and provide real-time tool-call feedback via CopilotKit
- Orchestrated the **Google Agent Development Kit (ADK)** for complex planning and RAG retrieval, integrating session management and MCP Server tool calls to dynamically route queries between live capabilities and internal knowledge
- Engineered a robust RAG pipeline leveraging **LangChain and Pinecone** to perform document embedding and indexing, ensuring **LLM** generation is grounded in precise source contexts
- Integrated web search and external tools through the **Model Context Protocol (MCP)**, streaming ADK events to the frontend for transparent and human-approvable agent execution
- Deployed on **AWS** with **ECS Fargate**, **ALB**, and **Amplify**; automated releases to **ECR** via GitHub Actions with **OIDC**

Post-Trade Allocation Platform Sep 2025 – Dec 2025

- Developed a scalable, **distributed microservices system** for allocation message recording, corrections, and cancellations with Spring Boot, Spring Cloud, and Kafka
- Designed and implemented the **Allocation Engine** using Spring Data JPA, QuickFIX/J, and Spring Security, capable of processing high-throughput FIX messages for execution validation and structured reporting
- Designed the **Kafka** architecture with **3 topics (allocation, correction, exception reporting) × 4 partitions each**, keyed by trade ID to preserve per-trade ordering; tuned acks and ISR for durability
- Secured APIs with **OAuth 2.0** and **Spring Cloud Gateway**, and deployed services with **Docker, Kubernetes**, and **Jenkins** for full CI/CD automation and dynamic scaling

TECHNICAL SKILLS

Programming Languages: Java, Python, TypeScript, JavaScript, SQL, C/C++
Backend & APIs: Spring Boot, Spring Cloud, Spring Data JPA, Spring Security, REST APIs, QuickFIX/J, OAuth 2.0
Data & Messaging: MySQL, PostgreSQL, Redis, Kafka, Pinecone
Cloud & DevOps: AWS, Docker, Kubernetes, Jenkins, GitHub Actions, Git
AI & Tools: RAG, LangChain, Google ADK, MCP, Claude Code, Codex, CopilotKit, Next.js, React