

Ruoke Zhang

✉ jaade1107@gmail.com

🌐 ruokezhang.github.io/en

📄 Ruoke Zhang

📞 206-699-2362

EDUCATION

University of Washington, Seattle

MS in Electrical & Computer Engineering

Sep 2025 – Mar 2027

The Ohio State University

BS in Computer Science; BS in Mathematics— GPA: 3.987/4.0 (Top 1%)

Aug 2021 – May 2024

EXPERIENCE

University of Washington

NeuroAI Lab Research Intern (Spatially-Aware Audio-Visual-Language Encoder)

Nov 2025 – Present

- Designed pipeline converting 360° video and FOA audio into HRTF-aligned NFOV binaural clips via spherical energy localization.
- Built Gemini 2.5 annotation pipeline selecting primary views and generating spatially grounded audio-visual event labels.
- Merged cross-view objects and applied spatial augmentations, expanding dataset to 200k spatial audio-visual objects.
- Built multimodal filtering using VGGish clustering, ImageBind similarity, and CLAP embeddings for aligned audio-visual events.

TikTok

Backend Software Engineer Intern

Apr 2025 – Aug 2025

Livestream Highlight Auto-Generation System

- Built a Python-Go auto-highlight service, orchestrated by RocketMQ tasks and Redis state recovery.
- Redesigned three-stage pipeline detecting Z-score peaks for PV, engagement, orders; aligning ASR timestamps; then running four LLM passes (GPT + Gemini) to extract, filter, refine clips, boosting GMV by 3.6 % in A/B testing.

Shanghai Artificial Intelligence Laboratory

Machine Learning Engineer Intern

Aug 2024 – Feb 2025

- Optimized GRScene dataset in Nvidia Isaac Sim, enhancing object properties and reducing GPU peak memory by 32.7%
- Conducted distributed communication tests using NCCL-Test and PyTorch Distributed across NVLink and TCP.
- Compared NCCL, Gloo, and RLLib to improve distributed training performance on multi-node, multi-GPU clusters.

PROJECTS

IntelliFlow - RAG-powered agent for System Design QA and blog publishing

Dec 2025

- Enabled multi-format document ingestion via Apache Tika and implemented vector storage using PostgreSQL pgvector
- Designed hybrid memory using Redis for real-time session buffering and PostgreSQL for long-term summaries memory.
- Achieved 95% Top-3 accuracy by implementing structure-aware chunking, Cohere ReRank, and Query Rewrite agents.
- Built an intent-driven MCP agent to autonomously orchestrate blog publishing to local files or GitHub repositories.

Preferred Time - Movie Recommendation Infrastructure

UW Hackathon '25

- Ingested high-volume user liking events via Kafka into S3 to support scalable offline retraining.
- Trained and deployed models using AWS SageMaker with DynamoDB-backed feature retrieval for inference.
- Instrumented Prometheus, Grafana, OpenTelemetry, and Jaeger to monitor inference latency and throughput.

Distributed KV Store

Course Project

- Implemented exactly-once RPC for a key-value store over lossy networks using retries and deduplication.
- Built Multi-Paxos state machine replication with ballot-based leader election, no-op hole-filling, and heartbeat-driven GC.
- Extended to a sharded key-value store across multiple Paxos replica groups, with a Paxos-replicated ShardMaster for configuration management and atomic shard migration carrying full KV and deduplication state.

RPC-Java

Dec 2025

- Built a Netty RPC framework with custom protocol, heartbeats, and Kryo/Protobuf/Hessian serialization.
- Ensured reliability using circuit breaker, Guava-based retries, and server-side token-bucket rate limiting.
- Integrated Nacos service discovery with client-side load balancing: weighted round-robin and least-connection.

SKILLS

Languages: Python(Advanced), Java(Advanced),Golang(Proficient), JavaScript, C++, HTML/CSS, PySpark

Frameworks: Spring Boot(Advanced), Flask(Advanced), React, Next.js, Node.js, FastAPI, Tailwind, Angular, Express.js, .NET

Cloud & Tools: AWS(Proficient), Azure, Google Cloud, Docker, Kubernetes, Terraform, Jenkins, Git

Databases: MySQL, PostgreSQL, MongoDB, Azure SQL Database, Azure Cosmos DB, SQL Server, GraphQL

Concepts: Backend Development, Distributed Systems, Microservices, REST APIs, Agile, System Design, Cloud Architecture