# Ran Li

Singapore | e1504040@u.nus.edu | (+65) 9081 2289 | Linkedin | GitHub

## CAREER OBJECTIVES

NUS Master of Computing student focusing on **backend development**, proficient in **Golang** and open to diverse technology stacks. Experienced in applying AI/LLM integration in real-world systems. Graduating in Dec 2026 and seeking an **internship** starting from Dec 2025 onwards, with a flexible end date.

**Strengths**: Passionate and analytical, with a strong scientific foundation and a relentless drive for growth, bringing a can-do attitude, rapid learning capability, and strong teamwork ability.

## EDUCATION

**National University of Singapore, Master of Computer Science** **Jan. 2025 – Dec. 2026**

- GPA: 4.75/5.0
- Coursework: Data Structures & Algorithms, Operating System, Advanced Computer Networks, Database Systems, Software Engineering, Artificial Intelligence, Neural Networks and Deep Learning, Big Data Systems for Data Science

**Southeast University, Bachelor of Engineering** **Aug. 2009 – June 2014**

## PROJECT EXPERIENCE

**Full Stack Project – Generation Blog** **Mar. 2025 – July 2025**

Links: Live Demo | GitHub

Tech Stack: *Go (Gin, Gorm), MySQL, Redis, Elasticsearch, OpenAI API, JWT, Nginx, Docker, ECharts, AWS, Vue, arco-design*

- Developed and deployed a **full-featured blogging platform** with conventional features such as user authentication, blog publishing, full-text search, and user engagement, integrated with **large language models** to support daily automated ArXiv paper aggregation, scoring, filtering, and publishing, as well as AI-assisted search and AI-powered writing.
- Designed the **system architecture**. Engineered RESTful backend APIs using **Gin**, implemented **JWT** authentication, **Redis** caching mechanisms, and **Elasticsearch**-driven full-text search. Managed relational data storage via **MySQL**, handling user records, file metadata, and system logs. Streamlined deployment processes through **Docker** containerization and AWS hosting with automated CI/CD pipelines.
- Created an **automated academic paper aggregation pipeline** from ArXiv, leveraging a dual-randomized scoring algorithm powered by large language models (LLMs) to evaluate research papers across innovation, depth, and practical relevance, generating **daily AI-driven analytical reports** of top-ranked papers.
- Implemented a **semantic search infrastructure** utilizing Elasticsearch, achieving millisecond-level granular search capabilities. Integrated LLM-driven keyword extraction to facilitate precise **natural language search** and context-sensitive interactions.
- **Optimized database performance** with a primary-replica MySQL architecture, introducing read-write separation combined with Redis caching, resulting in a **70% throughput increase** and consistently low-latency responses under high concurrency.

**Full Stack Project – AI Dialog Management System** **Feb. 2025 – Aug. 2025**

Links: Live Demo | GitHub

Tech Stack: *Go (Gin, Gorm, Urfave), MySQL/SQLite, Qdrant, Redis, OpenAI API, Context management*

- Architected a **multi-branch, tree structured** AI dialog management system based on a three-layer hierarchy (Session - Dialog - Conversation), where each interaction is represented as a node enriched with AI-generated summaries for **fast navigation**. Supports **dialog branching** from any node, **structured context reconstruction, semantic memory retrieval**, and annotation/bookmarking capabilities to streamline dialog management.
- Built backend using **Gin** framework with RESTful API design, supporting both **Web** (Vue) and **CLI** interfaces. Employed **MySQL** as transactional data store, **Qdrant** for efficient semantic vector search, and **Redis** caching with a 4-hour TTL strategy.
- Designed a **context-aware retrieval pipeline** that dynamically reconstructs conversation context by aggregating recent dialog history from Redis/MySQL and augmenting it with semantically similar dialogs from Qdrant (similarity $\geqslant 0.75$), enabling **accurate and coherent multi-turn responses** through context-rich dialog assembly.

**Backend Project – Distributed Cache System** **July 2025 – Sept. 2025**

Tech Stack: *Go, gRPC, Protocol Buffers, etcd, Consistent Hashing, LRU/LRU-2/LFU/FIFO, SingleFlight, Goroutines, Channels*

- Developed a highly available distributed caching system supporting **gRPC-based** communication between nodes, service registration and discovery, dynamic node management, and fast network topology convergence.
- Implemented **LRU**, **LRU-2**, **LFU**, and **FIFO** cache eviction strategies to handle resource limitations, along with a TTL mechanism using goroutines for automatic cache expiration.
- Optimized inter-node communication via gRPC/HTTP with **Protobuf** encoding and used consistent hashing for load balancing, while leveraging **SingleFlight** to prevent cache stampede and enhance performance.
- Configured a highly available **etcd cluster** for service registration and discovery, implemented dynamic node management using a **watch mechanism**, and ensured **concurrency safety** with channels for goroutine communication.

**Machine Learning Project - Object Detection Model on Aerial Imagery** **Mar. 2025 – Apr. 2025**

Tech Stack: *PyTorch, Ultralytics, YOLOv11, scikit-optimize, OpenCV, Matplotlib*

- Focused on **multi-class aerial image recognition** by fine-tuning a YOLOv11n model on the DOTA dataset, powering accurate detection of objects such as vehicles, planes, and ships in satellite imagery, achieving mAP50 = 0.653 and mAP50–95 = 0.403 on the test set.

- Built a complete training pipeline with **data augmentation**, **model reinitialization**, and **results tracking** using PyTorch and Ultralytics.
- Delivered significant reductions in tuning time and computational overhead by adopting **Bayesian Optimization** over traditional search strategies, accelerating convergence while minimizing resource-heavy retraining cycles, with over 50% improvement in search efficiency.
- Strong team player, able to communicate effectively within our diverse (China, Singapore, India), cross-functional team to deliver the project successfully.

## PROFESSIONAL EXPERIENCE

**Parametric Landscape Architect / Architecture & Engineers Co., LTD. of Southeast University** **July 2014 – Dec. 2024**

- Led the design of 20+ projects, leveraging computational design methods—such as parametric modeling, big data simulation and AI-assisted rendering—to enhance design accuracy, efficiency, and cost-effectiveness. Several projects received national or ministerial-level awards.
- Led teams of 2–6 members and served as the main liaison with other departments, strengthening leadership, cross-functional communication, and resilience under tight deadlines.
- Coordinated with diverse stakeholders (clients, consultants, government bodies), honing negotiation and stakeholder management skills to ensure alignment and successful project delivery.
- Mentored junior colleagues and promoted knowledge sharing, improving team efficiency, and fostering a collaborative, growth-oriented culture.

## HONORS & AWARDS

**National, Provincial or Municipal Outstanding Design Awards** **2015 - 2024**

- Awarded **three First Prizes** in provincial engineering and landscape design competitions, along with 10+ national and regional honors.

## SKILLSETS

**Programming:** Golang, Python, Java, SQL, JavaScript, NoSQL
**Tech Stack:** MySQL, Redis, Gin, GORM, gRPC, Linux, Bash, Docker, Kubernetes, Git, Postman, MongoDB, HTML, CSS, Node.js, Hadoop, Spark, PyTorch
**Language:** English (fluent, TOEFL: 114/120), Mandarin (native)