Juntao Chen

+86 18559163703 | juntaochen718@foxmail.com

Education

University of Melbourne, Melbourne, Australia 04/2026 –04/2030

Doctor of Philosophy (PhD) in Computer Science

Sun Yat-sen University (SYSU), Guangzhou, China 09/2021 –06/2025

Bachelor of Engineering in Software Engineering, GPA: **3.9/5.0**

IELTS: 7.5

Outstanding Student Scholarship, Sun Yat-sen University (2022)

Research Experience

What to Retrieve for Effective Retrieval-Augmented Code Generation? An Empirical Study and Beyond 07/2024 - Now

Wenchao Gu, ***Juntao Chen****, Yanlin Wang, Tianyue Jiang, Xingzhe Li, Mingwei Liu, Xilin Liu, Yuchi Ma, Zibin Zheng*

Will be published at ICSE 2026(CORE A*) **Topic**: AI for Software Engineering

arXiv:2503.20589 *https://arxiv.org/abs/2503.20589*

- **Led the end-to-end project implementation**, including initial literature review and method exploration, codebase development, experimental design, large-scale evaluations, thorough analysis of experimental results, and iterative refinement of the proposed methods. **Actively contributed** to paper writing, comprehensive performance assessment, and in-depth discussions of results and implications.
- Conducted a systematic **empirical study** on retrieval-augmented generation (RAG) for repository-level code generation, demonstrating that retrieving in-context code and relevant API documentation significantly improves LLM performance, while retrieving similar code can introduce noise and degrade results by up to 15%.
- Proposed **AllianceCoder**, a novel context-integrated framework that combines chain-of-thought prompting with semantic API retrieval for enhanced code generation.
- Achieved **state-of-the-art** performance on *CoderEval* and *RepoExec* benchmarks, improving Pass@1 by **10–20%** over strong baselines and achieving up to **46%** API recall, reducing the gap to oracle performance to within **5%**.

Professional Experience

Software Engineer Intern, *Lenovo Group Limited* 07/2024 - 10/2024

- **Led the full-stack development** of the NECAP Department Asset Management System. Developed the front-end with *Vue.js* and *Element Plus*, and built the back-end with *Golang* and the *GIN* framework. Implemented key features including asset inventory, asset transfer, and asset editing, ensuring smooth workflows and improving overall system efficiency for managing departmental assets.
- Implemented *WebSocket* for real-time requests and heartbeat detection, improving stability in high-concurrency environments. Deployed the system with *Nginx* for reverse proxy and load balancing to optimize request handling, enhance system resilience, and improve response speed by **20%** in high-concurrency scenarios.

Site Reliability Engineer Intern, *China Mobile Limited Fujian* 01/2024 - 02/2024

- **Developed an automated operations monitoring and fault diagnosis system** using *Python*, incorporating email alerts via the *smtplib* package and APIs for seamless automatic work order creation.
- **Built a large language model (LLM) toolchain** with *LangChain* to assist the operations team. Created various APIs with *Python* and the *Flask* framework for real-time server status updates, log retrieval, and critical data queries, which were integrated into the *LangChain* toolchain to improve the precision of LLM-generated outputs.

Software Engineer Intern, *Evecom Technology Co., Ltd.* 07/2023 - 08/2023

- Developed the **back-end** of a Legal Knowledge Question Bank System using *Java*, the *Spring Boot* framework, and *MyBatis* for database interactions, implementing management operations for question bank data. Integrated *Redis* for data caching, reducing API response time by **50%**.
- Implemented fine-grained access control using *Spring Security* and *JWT* for different user roles, enhancing system security and data protection. Extracted and organized legal questions from Word documents using *Apache POI* and regular expressions, achieving a **98%** extraction success rate while effectively handling exceptions.

Database Developer Intern, *AsiaInfo Technologies Limited* 01/2023 - 02/2023

- Designed the *Oracle database* for an Infrastructure Management system, developing efficient database schemas and optimizing indexing strategies to improve query performance. Additionally, implemented robust backup and recovery plans to ensure data integrity and maintain system reliability.
- Configured *Oracle Data Guard* for disaster recovery and implemented a regular backup schedule with point-in-time recovery (PITR) to ensure data availability and integrity. Utilized *Oracle Enterprise* for database optimization, improving data retrieval speed of key communication traffic reports by **50%**.

Projects

Foundation of the Public Data Hub at SYSU github.com/Elendil3703/data_center

Full Stack Developer & Team Leader

- Led the team to build a comprehensive central data platform, enabling independent system operation with streamlined data access and control. The platform provides end-to-end data management and analytics.
- Developed back-end APIs using *Java* and the *Spring Boot* framework for permission management, allowing system administrators to access shared data tables while restricting modifications to their own data.
- **Data Search**: Utilized *FlinkCDC* to monitor real-time database changes, with data streamed via *Kafka* to *ElasticSearch* for full-text search, enabling reliable data transmission across multiple data source clusters.
- Designed the front-end interface using *Vue.js* combined with *Element UI*. Created diverse data visualizations through *Kibana* for data analysis.

Campus Food Delivery Platform github.com/Elendil3703/food_delivery_system

Golang Backend Developer

- Built a comprehensive campus food delivery platform website, including both customer and merchant interfaces, with **1,563 registered users and O(100) orders per day**.
- Developed **back-end API** using *Golang* and the *GIN* framework, implementing core business modules such as product and order management. Leveraged *Goroutines* and *channels* to efficiently handle high-concurrency requests, managing up to 10,000 requests per second (RPS) while ensuring robust performance through *MySQL' s transaction* mechanisms and *Go' s sync package* for synchronization.

Awards

Innovation Award, HSBC Software Hackathon June 2023

Third Prize, MathorCup Big Data Modeling Contest November 2023

Third Prize, China College Math Modeling Contest, Guangdong September 2022

Skills

Programming Languages: C++, Java, Golang, Python, HTML, CSS, JavaScript, SQL

Technologies: Linux, Git, SpringBoot, GIN, Vue.js, Pytorch, MySQL, Oracle