

Li Rongzhi

+65 98119828 li.rongzhi@u.nus.edu [linkedin.com/in/rongzhi-li](https://www.linkedin.com/in/rongzhi-li) github.com/li-rongzhi

EDUCATION

National University of Singapore

Aug 2022 - Present

Bachelor of Computing in Computer Science (with Honours)

Singapore

- **GPA:** 4.83/5.0
- **Expected Graduation:** May 2026
- **Focus Area:** Artificial Intelligence, Database Systems
- Scheduled Exchange Program at **ETH Zurich, Switzerland** (AY2024/2025, Semester 1)
- **Relevant Coursework:** Foundation of AI and ML; Big Data Systems for Data Science; Database Systems Implementation

SKILLS

- **Programming Language:** Java, Python, JavaScript, TypeScript, Golang, C, SQL, Bash, HTML, CSS
- **Frameworks/Tools:** Numpy, Pandas, React.js, Node.js, Restful API, Django, Flask, TensorFlow, PyTorch, LangChain, AWS, Spark, Hadoop, MySQL, MongoDB, CI/CD pipelines, Gradle

HONOR

- **Awards:** Dean's List (Top 5%)
- **Scholarship:** ETH Zurich Exchange Student Scholarship; Science & Technology Undergraduate Scholarship

EXPERIENCE

MiniMax, Algorithm Engineer Intern

June 2024 - Present

- Enhanced the reporting of full-process metrics for the scheduling module, utilizing Grafana's visualization and anomaly alerting to achieve full lifecycle observability
- Optimized the data pipeline of the scheduling module, reducing each scheduling cycle time from 90 seconds to 50 seconds by merging and concurrently executing requests, significantly improving system performance
- Employed time series modeling to predict service loads based on features such as time and QPS, greatly reducing the frequency of load breaches, improving resource utilization efficiency while maintaining SLA compliance

National University of Singapore, Student Researcher

Sep 2023 - June 2024

- Researched on how to utilise large language models to simplify file processing, supervised by Prof. He Bingsheng
- Proposed a scalable ground truth dataset aimed at extracting content from financial reports, compiling a dataset of 100 publicly available 10-K forms
- Designed algorithms to accurately derive both structured and unstructured content from files in diverse formats by incorporating internal file structure analysis, OCR techniques and document layout analysis
- Developed a website - Legal Expert, facilitating smooth user engagement with LLMs leveraging knowledge base

Amaris.AI, Software Engineer Intern

Dec 2023 - Jan 2024

- Conducted in-depth research on extracting tabular data from PDF files; engineered a Camelot-based extractor surpassing standard tools in extraction precision, achieving an improvement of around 3.5%
- Created a prototype by integrating table extractor with Large Language Models, streamlining extraction and querying of tabular data in PDF files

PROJECTS

ProtoDB | Java, DBMS

Jan 2024 - Present

- A database system prototype implemented in Java, supporting SQL-like language parsing and CRUD operations
- Supported SQL-like parsing and database operations, including table creation, data query, update and deletion
- Designed database backup and recovery strategy, ensuring data reliability and efficient recovery after shutdown
- Adopted Two-Phase Locking (2PL) protocol, paired with Pessimistic Lock, to achieve correct serializability of schedules

pdf2table | Python, OpenCV

Jan 2024 - Present

- An open-source Python library designed to extract tables from PDF documents and convert them into analyzable data formats

- Employed advanced line detection and text block analysis algorithms based on OpenCV to identify table regions within PDFs, significantly outperforming deep learning models like DETR, with improvement in both accuracy and precision
- Integrated the TATR model for table structure recognition, preserving structural integrity and formatting of the extracted data
- Authored detailed documentation and usage examples, ensuring that developers can easily adopt and implement the library in various projects

Legal Expert | *JavaScript, React.js, Node.js, LLM*

Oct 2023 - Present

- An LLM-empowered website offering legal services, enriched with an extensive legal document knowledge base
- Built the frontend of the website with Ant Design components under React.js framework
- Assisted backend implementation, implemented several API under Node.js framework, including streaming output and legal document chunk retrieval