

# DONGJUN YU (余东骏)

@ E-mail: [ydsirhere@gmail.com](mailto:ydsirhere@gmail.com)

📖 Blog

🌐 YDSIR-NJU



## EDUCATION

Bachelor in Software Engineering

Software Institute, Nanjing University

📅 2019.9 – 2023.7

📍 Nanjing, China

– GPA: 4.42 / 5.0 (47/209, Top 23%)

– Distinguished Graduate of Nanjing University (南京大学优秀毕业生)

Master Candidate in Software Engineering (Research Program)

Software Institute, Nanjing University

📅 2023.9 – 2026.7

📍 Nanjing, China

Under the supervision of **Dr. Guoping Rong**

– GPA: 4.52 / 5.0 (University does not provide ranking)

## PUBLICATIONS

- Hao Zhang, Dongjun Yu, Lei Zhang, Guoping Rong, Yongda Yu, Haifeng Shen, He Zhang, Dong Shao and Hongyu Kuang. *AUCAD: Automated Construction of Alignment Dataset from Log-Related Issues for Enhancing LLM-based Log Generation*. (Internetwork' 25). (To appear, arXiv:2412.18835v2, CCF-C)
- Dongjun Yu, Xiao Yan, Zhenrui Li, Jipeng Xiao, Haochuan He, Yongda Yu, Hao Zhang, Guoping Rong, Xiaobo Huang. *SynthCoder: A Synthetical Strategy to Tune LLMs for Code Completion*. (arXiv:2508.15495v2, arXiv Preprint)

## WORK EXPERIENCE

LLM Development Engineer *Intern*

China Telecom Research Institute

SRDCloud Operating Center (研发云运营中心)

📅 2024.11 – 2025.11

📍 Guangzhou, China

Post-training **Seed-Coder-8B-Base** to serve China Telecom's code completion plugin *CodeFree*. Significantly improve performance in benchmarks and increases acceptance rate from 15% to 25%. Responsible for code implementation of training and evaluation.

Backend Development Engineer *Intern*

Meituan

Fundamental Development Platform (基础研发平台)

📅 2023.4 – 2023.8

📍 Shanghai, China

Leveraged explainable AI techniques to assess the contribution of each log template in anomaly detection algorithms and other metrics, a framework called *xLoad* (Available at <https://github.com/aioptionsplus/xload>) was validated in open-source datasets like **LogHub** and production environments, with nearly 10% of logs from **Raptor** (a critical observability and traceability infrastructure at Meituan) identified by *xLoad* was manually verified useless. Responsible for data preprocessing and application in production.

## RESEARCH INTEREST

Software Engineering

LLM4SE

AI Ops

Log

## LANGUAGE

- IELTS 7.5 (L: 8.5; R: 8; W: 6.5; S: 6)
- Fluent in spoken Cantonese

## AWARDS

- 教育部——华为“智能基座”未来之星  
2022

## COMPETITIONS

3rd "China Mobile Cloud Cup" Computing Power Network Application Innovation Competition

第三届“移动云杯”算力网络应用创新大赛

College Track - Grand Prize Ranked 1st

📅 2024.12

📍 Suzhou, China

Intelligent Code Review Based on Large Language Models — A comprehensive automatic code review solution based on **CarLLM** ( [doi.org/10.1145/3695993](https://doi.org/10.1145/3695993) )

integrated with **GitLab**.

Responsible for datasets construction, model testing and module integration.

2021-2022 17th Citi Financial Innovation Application Competition

第十七届“花旗杯”金融创新应用大赛

Grand Prize Ranked 1st

📅 2022.6

📍 Wuhan, China

TextualESG — An ESG evaluation platform based on multi-source textual information

Responsible for requirements analysis, process management, overall architecture and database design, CI/CD pipeline setup, deployment, and operations as well as some coding in both frontend and backend.