# DONGJUN YU (余东骏)

@ E-mail: ydjsirhere@gmail.com    % Blog    ⌂ YDJSIR-NJU    🎓 Publications

## EDUCATION

**Bachelor in Software Engineering**

**Software Institute, Nanjing University**

📅 2019.9 – 2023.7              📍 Nanjing, China

− **GPA: 4.38 / 5.0 (55/219，Top 25%)**

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

**Master's Degree Candidate in Software Engineering (Research-Oriented Program)**

**Software Institute, Nanjing University**

📅 2023.9 – 2026.7              📍 Nanjing, China

**Under the supervision of Prof. 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.* **(Internetware' 25).** (To appear, *CCF-C*)

## 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

Responsible for post-training **Qwen2.5Coder-7B-Instruct** to deliver the base LLM for *CodeFree*—the GitHub Copilot alternative of China Telecom — supporting code completion services. Significantly improve performance in **ExecRepoEval**, **CrossCodeEval** and other benchmarks.

**Backend Development Engineer** *Intern*

**Meituan**

*Fundamental Development Platform → Monitor System Group (基础研发平台 → 监控系统组)*

📅 2023.4 – 2023.7              📍 Shanghai, China

Responsible for controlling the volume of continuously generated log data from **Raptor**, a critical **observability and traceability infrastructure at Meituan**. 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/aiopsplus/xload`) was validated in **open-source datasets like LogHub and production environments**, where developers confirmed that nearly **10%** of logs identified by *xLoad* as less valuable could potentially be purged to raise efficiency in processing logs.

## RESEARCH INTEREST

Software Engineering    LLM4SE    AIOps    Log    Code Review

## LANGUAGE

- IELTS  7.5 (Listening: 8.5; Reading: 8; Writing: 6.5; Speaking: 6)
- CET-6  591
- CET-4  627
- *Fluent in spoken Cantonese*

## 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 (高校赛道 - 特等奖)** *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` *) integrated with GitLab.*

Responsible for datasets construction, model testing and module integration.

**2021-2022 17th Citi Financial Innovation Application Competition**

**Grand Prize (特等奖)** *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.