

# YUANKAI FAN

Researcher

 [github.com/Kaimary](https://github.com/Kaimary)

 +86 13681689643

 [kaimary1221@gmail.com](mailto:kaimary1221@gmail.com), [fanyuankai@fudan.edu.cn](mailto:fanyuankai@fudan.edu.cn)

 Institute of Artificial Intelligence, China Telecom (TeleAI), Shanghai, China



I am currently a Researcher at AI Infrastructure Research Center at TeleAI. I received my Ph.D in CS from FDU under the guidance of Prof. X. Sean Wang. My research lies at the intersection of AI infrastructure, machine learning systems, and database systems, with a particular focus on: 1) **Data & AI Infra**: multi-modal data processing & management and distributed caching systems; 2) **Efficient Training**: kernel fuse, dynamic parallelism; 3) **Efficient Inference**: caching, quantization and sparsity; 4) **Natural Language Interfaces**: nl2sql.

Before joining TeleAI, I was a postdoctoral research assistant at FDU (2023-2024), with a primary focus on database systems. Prior to that, I worked for Microsoft about seven years (2016-2023), and one year for DELL EMC<sup>2</sup> (2015-2016).

## EDUCATION

**PH.D. IN SOFTWARE ENGINEERING, FUDAN UNIVERSITY** 2017.9 – 2023.1  
*with Education Program Sponsor from Microsoft*

**B.S. IN COMPUTER SCIENCE&TECHNOLOGY, YANSHAN UNIVERSITY** 2009.9 – 2013.6

## RESEARCH PUBLICATIONS

### CONFERENCE

- 1 **Y. Fan**, et al, Grounding Natural Language to SQL Translation with Data-Based Self-Explanations. In International Conference on Data Engineering, 29-42. (**ICDE'25**)
- 2 T. Ren, K. Chen, **Y. Fan**, et al, The Power of Constraints in Natural Language to SQL Translation. In Proceedings of the VLDB Endowment, 18(7): 2097-2111. (**VLDB'25**)
- 3 **Y. Fan**, et al, MetaSQL: A Generate-then-Rank Framework for Natural Language to SQL Translation. In International Conference on Data Engineering, 1765-1778. (**ICDE'24**)
- 4 T. Ren, **Y. Fan**, et al, PURPLE: Making a Large Language Model a Better SQL Writer. In International Conference on Data Engineering, 15-28. (**ICDE'24**)
- 5 **Y. Fan**, et al, GAR: A Generate-and-Rank Approach for Natural Language to SQL Translation. In International Conference on Data Engineering, 110-122. (**ICDE'23**)
- 6 **Y. Fan**, et al, GenSQL: A Generative Natural Language Interface to Database Systems. In International Conference on Data Engineering, 3603-3606. (**ICDE'23**)
- 7 **Y. Fan**, et al, An Integrated Interactive Framework for Natural Language to SQL Translation. In International Conference on Web Information Systems Engineering, 643-658. (**WISE'23**)
- 8 **Y. Fan**, et al, GAR++: Natural Language to SQL Translation with Efficient Generate-and-Rank. In APWeb-WAIM International Joint Conference on Web and Big Data, 411-427 (**APWeb'24**)
- 9 P. Yu, **Y. Fan**, et al, Bucketing Is Not Enough: Efficient Text-to-Video Model Training Via Cascade-Based Spatio-Temporal Orchestration. In USENIX Symposium on Networked Systems Design and Implementation. (**NSDI'26**) [**Under Review**]

### JOURNAL

- 1 **Y. Fan**, et al, Rethinking Data in NL2SQL: A Survey of What We Have and What We Expect. In Vicinagearth. (**AIFlow'26**)
- 2 **Y. Fan**, et al, A Confidence-based Knowledge Integration Framework for Cross-Domain Table Question Answering. In Knowledge-based Systems, 306: 112718. (**KBS'24, IF 7.2**)
- 3 H. Chen, F. Ye, **Y. Fan**, et al, Staged Query Graph Generation based on Answer Type for Question Answering over Knowledge Base. In Knowledge-based Systems, 253: 109576. (**KBS'22, IF 8.8**)
- 4 T. Yu, Z. He, Z. Yang, F. Ye, **Y. Fan**, et al, Zebra: A Novel Method for Optimizing Text Classification Query in Overload Scenario. In World Wide Web, 26(3): 905-931. (**WWWJ'22, IF 3.7**)

### MISC.

- 1 **Y. Fan**, Z. He, X. Wang, the New Trend in Natural Language Interfaces for Databases - Opportunities and Challenges in the Era of Large Language Models. *Communications of the CCF*. (**CCCF-2023**)

## INDUSTRIAL EXPERIENCES

---

### **FUDAN UNIVERSITY (SHANGHAI, CHINA)**

2023.1 – 2024.11

#### *Postdoctoral Research Fellow*

- Research interests lie in artificial intelligence for database systems (AI4DB), data governance, and NLP related techniques.

### **MICROSOFT (SHANGHAI, CHINA)**

2016.5 – 2023.2

#### *Support Escalation Engineer*

- Act as an exclusive full-stack engineer in the team. My expertise encompass *Windows drivers, Windows Hardware Lab Kit, Azure IoT/Cognitive Services/Machine Learning Service*.

### **DELL EMC (SHANGHAI, CHINA)**

2015.6 – 2016.5

#### *Software Engineer*

- Develop the cloud-native distributed file system for Atmos storage product.

### **SHUHUI SYSTEM TECHNOLOGY CO., LTD. (SHANGHAI, CHINA)**

2013.7 – 2015.5

#### *Software Engineer*

- Develop add-ins (C++ program) for AutoCAD software.

## MISCELLANEOUS EXPERIENCES

---

### **HONORS AND AWARDS**

- Outstanding Paper Award of the 8<sup>th</sup> China Conference on Intelligent Technology and Big Data **2023**
- Second Prize in the SPDB-Baidu Geek Algorithm Challenge, NL2SQLT ask **2020**
- Student Scholarship Third Prize **2017**

### **ACADEMIC ACTIVITIES**

- Reviewer for journals (e.g., TKDE, Scientific Reports, IEEE Access, International Journal of Human-Computer Interaction, and Journal of Intelligent Information Systems), and conferences (e.g., DASFAA and ADMA). **2024-2025**