

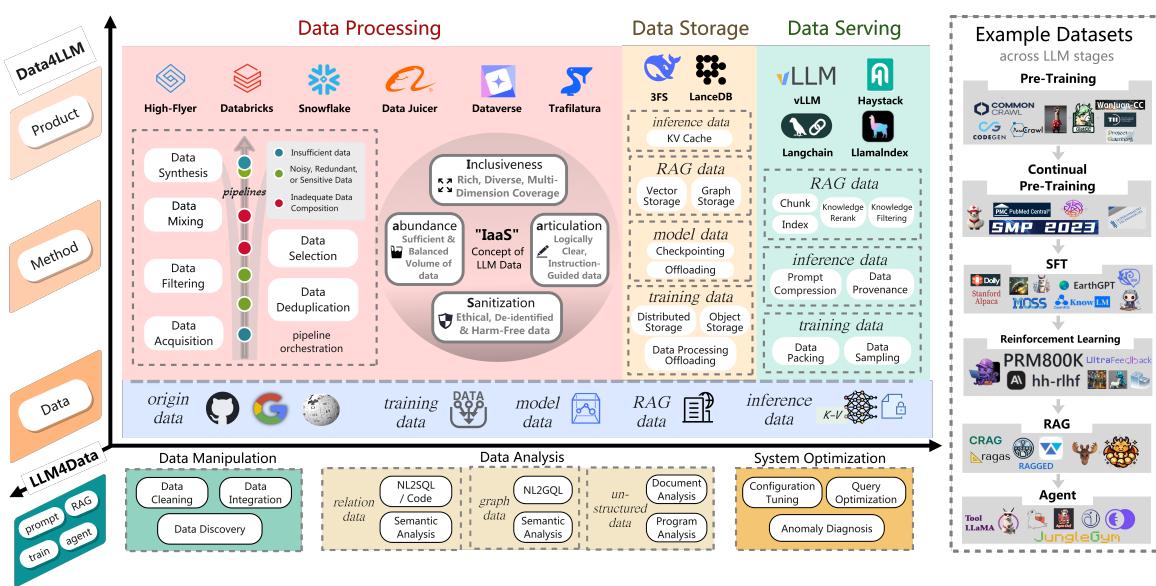
Xuanhe Zhou (周煊赫)

SJTU, Tenure-Track Assistant Professor, Faculty of Zhiyuan Honors Program

<https://db.zhouxh.store>

Research Interests

I focus on building AI-era data infrastructures (e.g., OpenMLDB). My work also supports scientific analytics (e.g., ST-Raptor) and underlying system maintenance (e.g., DB-GPT), optimization (e.g., LearnedRewrite), and migration (e.g., CrackSQL). I have published papers in data management (SIGMOD, VLDB, TKDE) and AI (NeurIPS) conferences and journals, with 2000+ citations on Google Scholar. I also maintain the most-starred [LLMxData](#) paper repository.



Representative Projects

As Project Leader

- [OpenMLDB](#) (1.6k stars): The Only MLOps solution from China in Gartner 2024; Utilized in **100+** applications.
- [DB-GPT](#) (647 stars): Asynchronous Multi-Agent Diagnosis; Selected as **Top 100** Open Achievements by BenchCouncil.
- [openGauss \(DBMind\)](#) (3k+ stars): Autonomous Database System with **7000+** contributors and **3.6 million** downloads.

As Major Contributor

- [BMTools](#) (2.8k stars): Tool Learning for LLMs; Open-Source Solutions of ChatGPT-Plugins
- [Bird-SQL](#) (1.4k stars): First cross-domain NL2SQL benchmark; Adopted by OpenAI for testing their finetuning service.

Awards

1. 2025 - ACM Jim Gray Doctoral Dissertation Honorable Mention Award (CCF A, first winner in China)
2. 2023 - VLDB Best Industry Paper Runnerup Award (CCF A, first author)
3. 2023 - BenchCouncil Top-100 Open Source Achievement (First Author)
4. 2024 - CCF Doctoral Dissertation Award (10 recipients in China)
5. 2022 - MSRA Fellowship (Top-12 in Asia)
6. 2022 - ByteDance Fellowship, 2022 (Top-10 in China)
7. 2022 - Tsinghua Top Grade Scholarship (清华特奖)
8. 2021 - Zhongshimo Fellowship (Highest Student Honor in Tsinghua CS)

Selected Publications

1. **OpenMLDB: A Real-Time Relational Data Feature Computation System for Online ML.**
Xuanhe Zhou, Wei Zhou, Bingsheng He, Mian Lu, Guoliang Li, Fan Wu, Yuqiang Chen.
SIGMOD (2025). [[paper](#)] [[code](#)]
2. **Cracking SQL Barriers: An LLM-based Dialect Translation System.**
Wei Zhou, Yuyang Gao, **Xuanhe Zhou**, Guoliang Li.
SIGMOD (2025). [[paper](#)] [[code](#)]
3. **D-Bot: An LLM-Powered DBA Copilot.**
Zhaoyan Sun, **Xuanhe Zhou**, Jianming Wu, Wei Zhou, Guoliang Li.
SIGMOD Demo (2025). [[paper](#)] [[code](#)]
4. **Revolutionizing Database QA with LLMs: Comprehensive Benchmark and Evaluation.**
Yihang Zheng, Bo Li, Zhenghao Lin, Yi Luo, **Xuanhe Zhou**, Chen Lin, Jinsong Su, Guoliang Li.
KDD (2025). [[paper](#)] [[code](#)]
5. **MemQ: Graph-Based Query Memory Prediction for Effective Workload Scheduling.**
Yang Wu, **Xuanhe Zhou**, Xiaoguang Li, Feifei Li, Yong Zhang.
ICDE (2025). [[paper](#)] [[code](#)]
6. **Robustness of Updatable Learning-based Index Advisors against Poisoning Attack.**
Yihang Zheng, Chen Lin, Xian Lyu, **Xuanhe Zhou**, Guoliang Li.
SIGMOD Full Paper, 2(1): V2mod010:1-V2mod010:26 (2024). [[paper](#)] [[code](#)]
7. **D-Bot: Database Diagnosis System using Large Language Models.**
Xuanhe Zhou, Guoliang Li, Zhaoyan Sun, Zhiyuan Liu, Weize Chen, et al.
VLDB Full Research Paper (2024). [[paper](#)] [[code](#)] **Benchcouncil Top 100 Open Achievements**
<http://dbgpt.dbmind.cn/dashboard>, with over 400 GitHub stars.
8. **Breaking It Down: An In-depth Study of Index Advisors.**
Wei Zhou, Chen Lin, **Xuanhe Zhou**, Guoliang Li.
VLDB Full Research Paper, (2024). [[paper](#)] [[code](#)]
9. **Chat2Data: An Interactive Data Analysis System with RAG, Vector Databases and LLMs.**
Xinyang Zhao, **Xuanhe Zhou**, Guoliang Li.
VLDB Demo, (2024). [[paper](#)] [[code](#)]
10. **LLM for Data Management.**
Guoliang Li, **Xuanhe Zhou**, Xinyang Zhao.
VLDB Tutorial, (2024). [[paper](#)] [[code](#)]
11. **TRAP: Tailored Robustness Assessment for Index Advisors via Adversarial Perturbation.**
Wei Zhou, Chen Lin, **Xuanhe Zhou**, Guoliang Li.
ICDE Full Paper, (2024).
12. **Automatic Index Tuning: A Survey.**
Yang Wu, **Xuanhe Zhou**, Yong Zhang, Guoliang Li.
TKDE, (2024).

13. Tool Learning with Foundation Models.

Yujia Qin, Shengding Hu, Yankai Lin, Weize Chen, Ning Ding, Ganqu Cui, Zheni Zeng, **Xuanhe Zhou**, Yufei Huang, et al.
CSUR, (2024).

14. DB-GPT: Large Language Model Meets Database.

Xuanhe Zhou, Zhaoyan Sun, Guoliang Li.
Data Science and Engineering, 9(1): 102-111 (2024). [[paper](#)] [[code](#)]

15. Grep: A Graph Learning Based Database Partitioning System.

Xuanhe Zhou, Guoliang Li, Wei Guo, Luyang Liu.
SIGMOD Full Paper, 1(1): 94:1-94:24 (2023). [[paper](#)] [[code](#)]

16. FEBench: A Benchmark for Real-Time Relational Data Feature Extraction.

Xuanhe Zhou, Cheng Chen, Kunyi Li, Bingsheng He, Mian Lu, Qiaosheng Liu, Wei Huang, et al.
VLDB Full Paper, 16(12): 3597-3609 (2023). [[paper](#)] [[code](#)] **Best Industry Paper Runnerup Award**

17. A Learned Query Rewrite System.

Xuanhe Zhou, Guoliang Li, Jianming Wu, Jiesi Liu, Zhaoyan Sun, Xinning Zhang.
VLDB Demo, 16(12): 4110-4113 (2023). [[paper](#)] [[code](#)]
http://rewrite_demo.dbmind.cn/

18. Learned index: A comprehensive experimental evaluation.

Zhaoyan Sun, **Xuanhe Zhou**, Guoliang Li.
VLDB Full Paper, 16(8): 1992-2004 (2023). [[paper](#)] [[code](#)]

19. DBAugur: An Adversarial-based Trend Forecasting System for Diversified Workloads.

Yuanning Gao, Xiuqi Huang, **Xuanhe Zhou**, Xiaofeng Gao, Guoliang Li, Guihai Chen.
ICDE Full Paper, 27-39 (2023). [[paper](#)] [[code](#)]

20. Can LLM Already Serve as A Database Interface? A Big Bench for Large-Scale Database Grounded Text-to-SQLs.

Jinyang Li, Binyuan Hui, Ge Qu, Jiaxi Yang, Binhu Li, Bowen Li, Bailin Wang, Bowen Qin, Ruiying Geng, Nan Huo, **Xuanhe Zhou**, Ma Chenhao, Guoliang Li, Kevin Chang, Fei Huang, et al.
NeurIPS Full Paper, (2024). [[paper](#)] [[code](#)]

21. Automatic Database Knob Tuning: A Survey.

Xinyang Zhao, **Xuanhe Zhou** (co-first author), Guoliang Li.
TKDE Survey Paper, 35(12): 12470-12490 (2023). [[paper](#)]

22. LearnedSQLGen: Constraint-aware SQL Generation using Reinforcement Learning.

Lixi Zhang, Chengliang Chai, **Xuanhe Zhou**, Guoliang Li.
SIGMOD Full Paper, 945-958 (2022). [[paper](#)] [[code](#)]

23. A Learned Query Rewrite System using Monte Carlo Tree Search.

Xuanhe Zhou, Guoliang Li, Chengliang Chai, Jianhua Feng.
VLDB Full Paper, 15(1): 46-58 (2021). [[paper](#)] [[code](#)]

24. AutoIndex: An Incremental Index Management System for Dynamic Workloads.

Xuanhe Zhou, Luyang Liu, Wenbo Li, et al.

ICDE Full Research, 2196-2208 (2022). [[paper](#)] [[code](#)]

25. Adaptive Code Learning for Spark Configuration Tuning.

Chen Lin, Junqing Zhuang, Jiadong Feng, Hui Li, **Xuanhe Zhou**, Guoliang Li.

ICDE Full Paper, 1995-2007 (2022). [[paper](#)] [[code](#)]

26. Machine learning for data management: A system view.

Guoliang Li, **Xuanhe Zhou**.

ICDE Tutorial, 3198-3201 (2022). (tutorial track)

27. openGauss: An Autonomous Database System.

Guoliang Li, **Xuanhe Zhou**, Ji Sun, Xiang Yu, et al.

VLDB Full Paper, 14(12): 3028-3041 (2021). [[paper](#)] [[code](#)]

28. DBMind: A Self-Driving Platform in openGauss.

Xuanhe Zhou, Lianyuan Jin, Ji Sun, Xinyang Zhao, Xiang Yu, Jianhua Feng, Shifu Li, Tianqing Wang, Kun Li, Luyang Liu.

VLDB Demo, 14(12): 2743-2746 (2021). [[paper](#)]

29. AI Meets Database: AI4DB and DB4AI.

Guoliang Li, **Xuanhe Zhou**, Lei Cao.

SIGMOD Tutorial, 2859-2866 (2021). (tutorial track)

30. Machine Learning for Databases.

Guoliang Li, **Xuanhe Zhou**, Lei Cao.

VLDB Tutorial, (2021). (tutorial track)

31. Database Meets Artificial Intelligence: A Survey.

Xuanhe Zhou, Chengliang Chai, Guoliang Li, Ji Sun.

TKDE Survey Paper, 1096-1116 (2020). [[paper](#)]

32. Query Performance Prediction for Concurrent Queries using Graph Embedding.

Xuanhe Zhou, Ji Sun, Guoliang Li, Jianhua Feng.

VLDB Full Paper, 13(9): 1416-1428 (2020). [[paper](#)] [[code](#)]

33. QTune: Query-Aware Database Tuning System with Deep Reinforcement Learning.

Guoliang Li, **Xuanhe Zhou**, Shifu Li, Bo Gao.

VLDB Full Paper, 12(12): 2118-2130 (2019). [[paper](#)] [[code](#)]

Last 5 Years' Most Cited VLDB Papers

34. Xuanyuan: An AI-Native Database.

Guoliang Li, **Xuanhe Zhou**, Sihao Li.

IEEE Data Engineering Bulletin, (2019). [[paper](#)]

Teaching

- **Data Structures, SJTU**, Teaching [Spring 2025]
- **Database Systems, SJTU**, Teaching [Spring 2025]