

Boyan LI

✉ bli303@connect.hkust-gz.edu.cn

🌐 liboyan.vip

🔗 github.com/BugMaker-Boyan

🔍 Google Scholar

Education

Hong Kong University of Science and Technology (Guangzhou)

Guangzhou, China

Ph.D. Student in Data Science and Analytics (DSA)

Sep. 2024 – Present

Supervisor: **Prof. Yuyu LUO**

Southern University of Science and Technology (SUSTech)

Shenzhen, China

B.Eng. in Computer Science and Technology

Sep. 2019 – Jul. 2023

GPA: 3.91/4.0 (Rank 2/183) | Honors: Outstanding Graduate, Highest Honors

Publications

* indicates equal contribution. **Boyan Li** is bolded.

First-Author Publications (Including Co-First)

- [1] **Alpha-SQL: Zero-Shot Text-to-SQL using Monte Carlo Tree Search**
Boyan Li, J. Zhang, J. Fan, Y. Xu, C. Chen, N. Tang, Y. Luo
International Conference on Machine Learning (ICML), 2025 (CCF-A)
Deployed in **Huawei Cloud DataArts Insight** and **Wuhan Urban Operation Agent**.
- [2] **DeepEye-SQL: A Software-Engineering-Inspired Text-to-SQL Framework**
Boyan Li, C. Chen, Z. Xue, Y. Mei, Y. Luo
ACM SIGMOD International Conference on Management of Data (SIGMOD), 2025 (CCF-A, Under Review)
- [3] **DeepEye: A Steerable Self-driving Data Agent System**
Boyan Li, Y. Peng, Y. Xie, S. Lu, Y. Zhu, X. Mu, X. Liu, Y. Luo
ACM SIGMOD International Conference on Management of Data (SIGMOD), 2026 (Demo, CCF-A, Under Review)
- [4] **DPC: Training-Free Text-to-SQL Candidate Selection via Dual-Paradigm Consistency**
Boyan Li*, O. O. K. Hei*, Y. Yu*, Y. Luo
Annual Meeting of the Association for Computational Linguistics (ACL), 2026 (CCF-A, Under Review)
- [5] **ROSE: An Intent-Centered Evaluation Metric for NL2SQL**
W. Pei*, S. Hou*, **Boyan Li***, H. Chen, Z. Shi, Y. Luo
Annual Meeting of the Association for Computational Linguistics (ACL), 2026 (CCF-A, Under Review)
- [6] **The Dawn of Natural Language to SQL: Are We Fully Ready?**
Boyan Li, Y. Luo, C. Chai, G. Li, N. Tang
Proceedings of the VLDB Endowment (VLDB), 2024 (CCF-A)
- [7] **DeepVIS: Bridging Natural Language and Data Visualization Through Step-wise Reasoning**
Z. Shuai*, **Boyan Li***, S. Yan, Y. Luo, W. Yang
IEEE VIS Conference (VIS), 2025 (CCF-A)

Co-authored Publications

- [8] **A Survey of NL2SQL with Large Language Models: Where are we, and where are we going?**
X. Liu, S. Shen, **Boyan Li**, P. Ma, R. Jiang, Y. Zhang, J. Fan, G. Li, N. Tang, Y. Luo
IEEE Transactions on Knowledge and Data Engineering (TKDE), 2025 (CCF-A)
- [9] **nvBench 2.0: A Benchmark for Natural Language to Visualization under Ambiguity**
T. Luo, C. Huang, L. Shen, **Boyan Li**, S. Shen, W. Zeng, N. Tang, Y. Luo
Conference on Neural Information Processing Systems (NeurIPS), 2025 (CCF-A)
- [10] **NL2SQL-BUGs: A Benchmark for Detecting Semantic Errors in NL2SQL Translation**
X. Liu, S. Shen, **Boyan Li**, N. Tang, Y. Luo
ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD), 2025 (CCF-A)
- [11] **Augmenting Realistic Charts with Virtual Overlays**
Y. Shi, **Boyan Li**, Y. Luo, L. Chen, N. Tang
CHI Conference on Human Factors in Computing Systems (CHI), 2025 (CCF-A)
- [12] **EllieSQL: Cost-Efficient Text-to-SQL with Complexity-Aware Routing**
Y. Zhu, R. Jiang, **Boyan Li**, N. Tang, Y. Luo
Conference on Language Modeling (COLM), 2025
- [13] **A Survey of Data Agents: Emerging Paradigm or Overstated Hype?**
Y. Zhu, ..., **Boyan Li**, et al.
IEEE Transactions on Knowledge and Data Engineering (TKDE), 2025 (Under Review)

- [14] **Advances and Challenges in Foundation Agents**
 B. Liu, ..., **Boyan Li**, et al.
arXiv preprint, 2025
[Ranked Top-1 in Hugging Face Daily Papers.](#)
- [15] **Are Large Language Models Good Statisticians?**
 Y. Zhu, S. Du, **Boyan Li**, Y. Luo, N. Tang
Conference on Neural Information Processing Systems (NeurIPS), 2024 (CCF-A)
- [16] **A Plug-and-Play Natural Language Rewriter for Natural Language to SQL**
 P. Ma, **Boyan Li**, R. Jiang, J. Fan, N. Tang, Y. Luo
arXiv preprint, 2024

Open Source & Systems

- **DeepEye: Visual Workflow Data Agent System**
 - Led the development of a low-code orchestration platform for transparent data analysis.
 - **Award:** “Best Open-source Project Award” at AI Agent 2025 Competition.
- **Alpha-SQL (ICML 2025)** [\[GitHub\]](#)
 - Official implementation of the Monte Carlo Tree Search based Text-to-SQL agent.
 - **Impact:** Deployed in **Huawei Cloud DataArts Insight** and **Wuhan Urban Operation Agent**.
- **NL2SQL Handbook (TKDE 2025)** [\[GitHub\]](#)
 - A comprehensive and continuously updated roadmap for NL2SQL researchers.
 - Widely recognized as a seminal resource for tracking state-of-the-art advancements.
- **NL2SQL360 Evaluation Framework (VLDB 2024)** [\[GitHub\]](#)
 - A comprehensive testbed for fine-grained evaluation of NL2SQL solutions.

Experience

Tsinghua University (Database Group)	<i>Beijing, China</i>
<i>Exchange Student (Mentor: Prof. Guoliang Li)</i>	<i>Jun. 2025 – Dec. 2025</i>
Huawei Central Research Institute	<i>Shenzhen, China</i>
<i>Research Intern</i>	<i>Jul. 2022 – Sep. 2022</i>

Honors & Awards

Research Excellence Awards, DSA Excellence Awards 2025	<i>Dec. 2025</i>
Runner-up, Best Project Award, DSAA 6000R, HKUST(GZ)	<i>Nov. 2025</i>
Best Open-source Project Award, AI Agent 2025 Competition	<i>Oct. 2025</i>
Merit Prize, DSA Excellent Research Award 2024	<i>Jan. 2025</i>
Best Project Award, DSAA 5021, HKUST(GZ)	<i>Dec. 2024</i>
Highest Honors in Computer Science & Engineering, SUSTech	<i>Jun. 2023</i>
Outstanding Graduate, SUSTech	<i>Jun. 2023</i>
Huawei Scholarship, Huawei Technologies	<i>May. 2022</i>
Excellent Student Scholarship, SUSTech	<i>2020, 2021, 2022</i>
National Encouragement Scholarship, Ministry of Education of China	<i>Apr. 2020</i>

Services

Conference Reviewer: International Conference on Learning Representations (ICLR)	<i>2026</i>
Invited Talk: Huawei Squirrel Webinar, “DeepEye-SQL”	<i>Jan. 2026</i>