

Yupeng Xie

Tel: +86-13129159589 | yxie740@connect.hkust-gz.edu.cn
No. 1 Yingcui North Road, Shunde District, Foshan, Guangdong, China 528315

EDUCATION

The Hong Kong University of Science and Technology (Guangzhou)

Ph.D. Student, Data Science and Analytics

Visiting Student, Information Hub - Data Science and Analytics Thrust

Guangzhou, China

Mar. 2025 - Present

Sep. 2023 - Feb. 2025

- Supervised by Assistant Professor Yuyu Luo.
- Research Interests:** Data Agents, Intelligent Data Visualization, and Data Storytelling.

Tsinghua University

Visiting Student, Department of Computer Science and Technology - Database Group

Beijing, China

Aug. 2022 - Sep. 2023

- Supervised by Professor Guoliang Li.

Qinghai University

Master of Science in Artificial Intelligence

Qinghai, China

Sep. 2021 - Jun. 2024

- Supervised by Professor Guoliang Li.

GPA: 89.47/100

PUBLICATIONS

[1] **Yupeng Xie**, Zhenyang Wang, Liangwei Wang, Jiayi Zhu, Zhouan Shen, Yuyu Luo. "DataMagic: Authoring Data Videos through Declarative Multi-Agent Orchestration." *IEEE Visualization and Visual Analytics (VIS)*, 2026 (Long Paper).

- **DataMagic.** An end-to-end system for authoring data videos from raw tabular data, featuring the declarative DVSpec and a Generate-then-Orchestrate multi-agent framework.

[2] **Yupeng Xie**, Chen Ma, Zhenyang Wang, Liangwei Wang, Jiayi Zhu, Chuxuan Zeng, Zhouan Shen, Boyan Li, Yuyu Luo. "DataMagic: Transforming Tabular Data into Data Insight Video." *Proceedings of the VLDB Endowment (Demo)*, 2026.

[3] **Yupeng Xie**, Zhiyang Zhang, Yifan Wu, Sirong Lu, Jiayi Zhang, Zhaoyang Yu, Jinlin Wang, Sirui Hong, Bang Liu, Chenglin Wu, Yuyu Luo. "VisJudge-Bench: Aesthetics and Quality Assessment of Visualizations." *International Conference on Learning Representations (ICLR)*, 2026.

- **VisJudge-Bench.** The first comprehensive benchmark for visualization aesthetics and quality assessment, with 3,090 expert-annotated samples across 32 chart types and a model designed to better align MLLM judgments with humans.

[4] Yinghao Tang, Xueding Liu, Boyuan Zhang, Tingfeng Lan, **Yupeng Xie**, Jiale Lao, Yiyao Wang, Haoxuan Li, Tingting Gao, Bo Pan, Luoxuan Weng, Xiuqi Huang, Minfeng Zhu, Yingchaojie Feng, Yuyu Luo, Wei Chen. "IGenBench: Benchmarking the Reliability of Text-to-Infographic Generation." *Annual Meeting of the Association for Computational Linguistics (ACL)*, 2026 (Long Paper).

[5] Boyan Li, Zhuowen Liang, **Yupeng Xie**, Xiaotian Lin, Tianqi Luo, Xinyu Liu, Yizhang Zhu, Zhangyang Peng, Yuan Li, Zhengxuan Zhang, Jiayi Zhang, Nan Tang, Guoliang Li, Yuyu Luo. "DataSpace: Benchmarking Data Agents for Verifiable Analytics over Heterogeneous Workspaces." *arXiv preprint arXiv:2608.03451*, 2026.

[6] Boyan Li, Yiran Peng, **Yupeng Xie**, Sirong Lu, Yizhang Zhu, Xing Mu, Xinyu Liu, Yuyu Luo. "DeepEye: A Steerable Self-driving Data Agent System." *ACM SIGMOD International Conference on Management of Data (Demo)*, 2026.

[7] Yinghao Tang, **Yupeng Xie**, Yingchaojie Feng, Tingfeng Lan, Wei Chen. "Demonstrating ViviDoc: Generating Interactive Documents through Human-Agent Collaboration." *Annual Meeting of the Association for Computational Linguistics (System Demonstrations)*, 2026.

[8] Yiyu Chen, Yifan Wu, Shuyu Shen, **Yupeng Xie**, Leixian Shen, Hui Xiong, Yuyu Luo. "ChartMark: A Structured Grammar for Chart Annotation." *IEEE Visualization and Visual Analytics (VIS)*, 2025 (Short Paper).

[9] **Yupeng Xie**, Yuyu Luo, Guoliang Li, Nan Tang. "HAICart: Human and AI Paired Visualization System." *Proceedings of the VLDB Endowment*, 2024, 17(11): 3178-3191.

- **HAICart.** A reinforcement-learning-based human-AI paired visualization system that combines automatic visualization generation with user-guided visualization hint selection.

[10] **Yupeng Xie**, Yuyu Luo, Jianhua Feng. "Navi: A Natural Language-powered Data Analytics System." *Journal of Software*, 2024, 35(3): 1194-1206.

[11] Yuyu Luo, Xuedi Qin, **Yupeng Xie**, Guoliang Li. "Intelligent Data Visualization Analysis Techniques: A Survey." *Journal of Software*, 2024, 35(1): 356-404.

ACADEMIC SERVICE

Committee Chair (Publicity & Data), KDD Cup 2026: Data Agents for Complex Data Analysis.

Program Committee Member, ADS 2026: The Joint Workshop on Agentic Data Systems and Data-Centric AI (co-located with VLDB 2026).

Reviewer, Conference on Neural Information Processing Systems (NeurIPS), 2026.

WORK EXPERIENCE

Guangdong Xingfu Century Group

Java Developer

Guangzhou, China

Sep. 2018 - May 2020

- Participated in the back-end development of the website, mobile app, and WeChat mini-program, independently developed by the company.

PROJECT EXPERIENCE

Company Project - Mingpinmao Shopping Mall

Jan. 2019 - Jul. 2019

Mingpinmao, a stylish and sophisticated shopping mall, curates high-quality lifestyle goods, offering a selection of premium products from around the globe. The mall is designed with a distributed architecture. Key systems include product, membership, order, finance, marketing, payment, account, supplier, profit-sharing, supplier settlement, and messaging systems.

- Backend microservices cluster built using SpringCloud technology stack.
- Implemented large-volume distributed file storage based on FastDFS.
- Used Redis for caching to reduce database load and improve query speed.

Customized Fitness Mall Platform Based on Microservices Architecture

Oct. 2018 - Mar. 2019

A personalized fitness shopping platform developed on a microservices architecture, intelligently tailoring content to users' characteristics, shopping habits, and browsing history to select products that suit their individual needs. This platform aids merchants in precisely targeting user groups, boosting browsing and conversion rates for effective targeted marketing.

- Utilized SpringCloud microservices architecture and Vue.js for frontend and backend separation development.
- Nginx used for reverse proxy, RabbitMQ for inter-service communication, and Elasticsearch as the search engine.
- Based on the design of a large e-commerce database to address the complex management of SPU and SKU.
- Used Thymeleaf for template generation and page static, improving view response speed and platform concurrency.
- Achievement: **Distinguished Undergraduate Thesis.**

Modern Museum Unmanned Guide System

Jun. 2018 - Dec. 2018

A cultural heritage protection project that merges IoT technology and microcontrollers. It features two modes: one for viewers and another for administrators. Viewers can scan a QR code to access a web-based voice guide assistant, interacting with cultural relics for historical information in multiple languages. Administrators can control the circuit system through voice and touch, analyzing exhibit data for strategic museum planning.

- The backend is developed using the SSM framework, facilitating communication with microcontrollers via Socket.
- Integrated iFLYTEK's SDK for voice interaction capabilities.
- Achievement: Won the third prize in the "CITIC Cup" IoT competition in Guangdong Province.

AWARDS

International & Regional

Silver Medal, 51st Geneva International Exhibition of Inventions, DeepEye Data Agent System, 2026

Silver Medal, 51st Geneva International Exhibition of Inventions, Low-Altitude Economy Ground-Air Collaborative Delivery Digital Twin Platform, 2026

First Place, Greater Bay Area Seven Universities Joint Graduate Academic Poster Competition, 2025

National Level

Best Open-source Project Award, AI Agent 2025 Competition

National Scholarship

National Encouragement Scholarship

China Telecom Scholarship · Fly Young Award (Top 1%)

Second Prize at the National Level and First Prize at the Provincial Level in the Lanqiao Cup National Programming Contest

Second Prize in the National College Students' Mobile Internet Application Development Innovation Competition

Second Prize at the National Level and Grand Prize at the Provincial Level in the National E-Business Innovation Competition

Provincial Level

Third Prize in the Guangdong Provincial College Students' Computer Design Competition

School Level

First Place in Men's 800m and 1500m at the School Sports Meet

SKILLS

Languages: CET-6, Mandarin: Native, Cantonese: Fluent.

Certificates: NCRE Level 2 - C Language, IBM Certified Senior Software Developer, National Advanced Certificate in Software Development (C/C++), C1 Driver's License.