

Zhifan Zhou

skyzhou@andrew.cmu.edu | (+1) 412-352-1112 | (+86) 189-5744-2769

EDUCATION

Carnegie Mellon University

B.S. in Information Systems; additional major in Artificial Intelligence

Pittsburgh, PA

Aug. 2025 – Present

- **Expected Graduation:** June 2029
- **Honors:** Dean's List, 2026
- **Selected Coursework:** Fundamentals of Programming & Computer Science, Concepts of Mathematics, Matrices and Linear Transformations, Concepts in Artificial Intelligence

PUBLICATIONS

DataMaster: Data-Centric Autonomous AI Research | *Under Review for NeurIPS 2026*

[arXiv: 2605.10906](#) [GitHub: DataMaster](#)

- Developed an autonomous framework for discovering, selecting, and cleaning external data to improve downstream machine learning and post-training performance.
- Designed a tree-structured search architecture with a shared data pool, enabling exploration of open-ended data spaces while preserving global memory across branches.
- Increased the MLE-Bench Lite medal rate by 32.27% and outperformed the instruct baseline on PostTrainBench GPQA by 0.67 percentage points.

RELEVANT EXPERIENCE

Research Intern

May 2026 – Present

NLP Lab, Eastern Institute of Technology

Ningbo, China

- Develop a streaming omni model for real-time multimodal interaction using a full-duplex training paradigm, under the supervision of [Prof. Xiaoyu Shen](#).
- Develop post-training data pipelines by curating omni-modal conversational datasets, aligning cross-modal signals, and defining fine-grained timestamp annotations for precise temporal supervision.

Research Intern

Feb. 2026 – Present

MAGIC Lab, Shanghai Jiao Tong University

Shanghai, China

- Conduct research on self-evolving agents and AI-for-science systems under [Prof. Siheng Chen](#).
- Contribute to projects spanning multi-agent governance, autonomous AI systems, and scientific discovery workflows.

Member

Sept. 2025 – Present

Data Science Club, Carnegie Mellon University

Pittsburgh, PA

- Apply PyTorch and deep learning techniques in hands-on workshops covering image classification and end-to-end model development.

TECHNICAL SKILLS

Programming Languages: Python, Java, C

Libraries & Frameworks: PyTorch, NumPy, pandas, Matplotlib

Languages: Mandarin Chinese (Native), English (Professional Working Proficiency)

Interests: Basketball, Table Tennis, Badminton, Tennis, Go (1 dan), Taekwondo (Red-Black Belt)