

ZHE YANG

+65 8264 4077 | [✉ zhe012@e.ntu.edu.sg](mailto:zhe012@e.ntu.edu.sg) | [🏠 Homepage](#)

Education

Nanyang Technological University Sep. 2025 – Present
Master of Science in Artificial Intelligence (MSAI) Singapore

- GPA: **5.0** / 5.0
- Advisor: Prof. Ziwei Liu

The Chinese University of Hong Kong Sep. 2021 – May 2025
B.Sc. in Computer Science (First Class Honor) Hong Kong SAR, China

- Major GPA: **3.802** / 4.000 (top 5%)
- Advisor: Prof. Michael R. Lyu

Peking University July 2023
Summer School (Grade: 94) Beijing, China

Experience

- **MMLab, Nanyang Technological University** Sep. 2025 – Present
Research Intern, advised by Prof. Ziwei Liu
Multimodal Agents, Contextual Benchmarking, Personalized File-System Memory
- **ARISE Lab, The Chinese University of Hong Kong** Jun. 2024 – May 2025
Research Intern, advised by Prof. Michael R. Lyu
LLM Memorization, Data Privacy, Tokenization Security
- **Geovis Technology Co., Ltd** Jul. 2024 – Aug. 2024
Algorithm Intern, Product Research & Development Department
Knowledge Graph, Model Fine-tuning, Neo4j, Automated Data Acquisition
- **Online Research Seminar** Feb. 2024 – Mar. 2024
Research Participant, advised by Prof. Dejan Marković
EEG Signal Processing, Disease Onset Prediction
- **IDEA Lab, The Chinese University of Hong Kong** Jun. 2023 – Sep. 2023
Research Intern, advised by Prof. Tsung-Yi Ho
Adversarial Attacks, Facial Recognition Robustness

Publications

- **HippoCamp: Benchmarking Contextual Agents on Personal Computers** **ECCV 2026**
Under Review
Zhe Yang, Shulin Tian, Kairui Hu, Shuai Liu, Hoang-Nhat Nguyen, Yichi Zhang, Zujin Guo, Mengying Yu, Zinan Zhang, Jingkang Yang, Chen Change Loy, and Ziwei Liu.
A benchmark for evaluating agents' capabilities on multimodal file management in user-centric environments to model individual user profiles and search from massive personal files for context-aware reasoning.

- FileGram: Grounding Agent Personalization in File-System Behavioral Traces** **ECCV 2026**
Under Review
 Shuai Liu, Shulin Tian, Kairui Hu, Yuhao Dong, **Zhe Yang**, Bo Li, Jingkang Yang, Chen Change Loy, and Ziwei Liu.
A framework grounding agent personalization in file-system behavioral traces with a scalable data engine and bottom-up memory architecture.
- Data Compressibility Quantifies LLM Memorization** **TMLR**
Under Review
 Yizhan Huang, **Zhe Yang**, Meifang Chen, Nianchen Huang, Jianping Zhang, and Michael R. Lyu.
A compressibility-based framework to quantify and characterize data memorization in LLMs.
- Understanding Secret Leakage Risks in Code LLMs: A Tokenization Perspective** **ACL 2025**
Findings
NeurIPS 2025
DL4C Workshop
 Meifang Chen, **Zhe Yang**, Nianchen Huang, Yizhan Huang, Yichen Li, Zihan Li, and Michael R. Lyu.
Revealed that BPE tokenization creates “Gibberish Bias” that increases secret memorization and leakage risks in Code LLMs.
- RAUIE: A Relation-Augmented Document-level Event Extraction Model Based on UIE** **AINIT 2024**
 Y. Liang, H. Jia, L. Wang, Y. Huang, and **Zhe Yang**.
Introduced entity relationships into prompts via prompt-based learning and designed a relation-augmentation algorithm to improve document-level event extraction on long-tail distributions.

Honors & Awards

- **HKSAR Government Scholarship** (Top 1% in Hong Kong) 2025
- **Silver Award for Outstanding Academic Performance** (Top 3%, CUHK) 2024
- **Dean’s List** (2022–2025, top 10% in the department), CUHK 2022–2025
- CUHK Student Envoy 2021

Skills

Programming: Python, C/C++, Java, HTML, R, MATLAB, RISC-V Assembly

Frameworks & Tools: PyTorch, MySQL, Neo4j, Linux, Git, L^AT_EX, Figma, Canva

Languages: Mandarin (Native), English (GRE: 322)