

Ziyang (Deric) Zhou

 <https://github.com/Azhou0> |  [azhou0.github.io](https://github.com/azhou0) |  ZiyangZhou
 Google Scholar |  azhou131416@gmail.com |  Ziyang.Zhou22@student.xjtlu.edu.cn

EDUCATION

- 2026 – Present **Nanyang Technological University (NTU)**, Singapore
Ph.D. Student in Computer Science, College of Computing and Data Science (CCDS)
Advisor: Prof. Yang Liu
- 2022 – 2026 **Xi'an Jiaotong–Liverpool University (XJTLU)**, Suzhou, China
B.Sc. Information and Computer Science (Cum. GPA: 3.7/4.0)
Selected Coursework: Linear Algebra, Calculus, Multivariable Calculus (4.0/4.0)

RESEARCH INTERESTS

- **Large Language Model Agents:** multi-agent system; agentic reasoning; agentic memory; self-evolving agents; agent reinforcement learning; retrieval-augmented generation
- **AI for Science & Health:** deep research; Human-centric AI; affective and cognitive computing
- **Audio & Natural Language Processing:** acoustic scene classification; emotion and empathy modeling; multimodal audio-text reasoning

PUBLICATIONS

RAM-SD: Retrieval-Augmented Multi-agent framework for Sarcasm Detection

Z. Zhou*, Z. Liu*, Y. Wang, Y. Lin, Y. Chen.

ACL 2026 Main Conference. [arXiv:2601.17002](https://arxiv.org/abs/2601.17002)

Following the TRACE: A Structured Path to Empathetic Response Generation with Multi-Agent Models

Z. Liu, Z. Zhou, Y. Li, H. Zhang, Y. Chen.

ICASSP 2026. [arXiv:2509.21849](https://arxiv.org/abs/2509.21849)

SEVADE: Self-Evolving Multi-Agent Analysis with Decoupled Evaluation for Hallucination-Resistant Irony Detection

Z. Liu*, Z. Zhou*, Y. Li, M. Hu, Y. Pan, Z. Xu, Y. Chen.

AAAI 2026. [arXiv:2508.06803](https://arxiv.org/abs/2508.06803)

Agent KB: Leveraging Cross-Domain Experience for Agentic Problem Solving

X. Tang*, T. Qin*, T. Peng*, Z. Zhou, D. Shao, T. Du, W. Zhou et al.

ICML 2025 Workshop on CFAgentic (Best Paper Runner-up Award). [arXiv:2507.06229](https://arxiv.org/abs/2507.06229)

CAF-I: A Collaborative Multi-Agent Framework for Enhanced Irony Detection

Z. Liu*, Z. Zhou*, M. Hu (*equal contribution).

ICML 2025 Workshop on CFAgentic (Oral Award). [arXiv:2506.08430](https://arxiv.org/abs/2506.08430)

AdapTF-SepNet: AudioSet-driven Adaptive Pre-training of TF-SepNet

Z. Zhou, Z. Yin, Y. Cai, S. Li, X. Shao.

DCASE 2025 Challenge Task 1, Technical Report.

EchoTwin-QA: A Dual-Tower BeatsBERT System for Audio Question Answering

Z. Yin, Z. Zhou, Y. Cai, S. Li, X. Shao.

DCASE 2025 Challenge Task 5, Technical Report.

WORK EXPERIENCE

LLM Researcher

EigenAI Oct 2025 – Feb 2026, Remote

- Contributed to **SEW-DC**, a data-centric RL framework that evolves robust long-horizon web agents through primitive-guided co-evolution.
- Led the development of the **Data Agent module**, generating slot-conditioned curriculum tasks across decision slots, target options, and difficulty levels.
- Designed validator-based filtering and supported MCPMark WebArena evaluation, improving the controllability of synthetic training data and robustness analysis.

Algorithm Engineer Intern

AxiomsTen Jun 2025 – Dec 2025, Shenzhen

- Implemented a lightweight cloud-based Voice Activity Detection (VAD) module for internal testing, integrating it into the speech pipeline alongside ASR and emotion recognition.
- Built FastAPI-based agent services that generate event-specific summaries and mindset extraction from user interactions, supporting adaptive and personalized responses.
- Collaborated with the algorithm team to design a Graph-RAG abstraction layer for representing and linking user events, enhancing the foundation for context reasoning and mindset modeling.
- Led the design and iteration of the agent workflow and prompt system, continuously refining multi-turn reasoning behavior to align with evolving product requirements.

LLM Agent Collaborator

OPPO Personal AI Mar 2025 – Jun 2025, Remote

- Helped design retrieval-augmented **agent memory** improving GAIA performance by **18.7%**.
- Curated and constructed the agent knowledge base by combining LLM-generated trajectories with human-guided reasoning analysis.
- Conducted extensive experiments across multi-agent benchmarks (GAIA, SWE-bench, etc.) with different agent systems (smolagent, Openhands, etc.) to evaluate cross-domain efficiency and generalization.

Automation Intern

Paslin China Jan 2025 – Mar 2025, Shanghai

- Observed the operational structure of a large-scale manufacturing enterprise, gaining a clear understanding of cross-departmental collaboration between design, engineering, and production teams.
- Learned to use EPLAN for creating electrical control diagrams and explored the logic of automated production lines in the automotive sector.
- Developed an initial interest in how industrial workflow design and system integration could inspire later research on intelligent automation and agentic decision systems.

SELECTED RESEARCH PROJECTS

Multi-Agent for Irony Detection & Empathy Response Generation Jun 2025 – Present Suzhou
XJTLU — Human-Centered Language Computing Group (Advisor: [Dr. Yangbin Chen](#))

- Designed and led three successive frameworks for irony detection — **CAF-I**, **SEVADE**, and **RAM-SD** — progressively improving multi-agent reasoning, evaluation reliability, and contextual understanding.
- In **CAF-I**, led the design of a collaborative multi-agent framework for irony detection, achieving a **+4.82%** Macro-F1 improvement, and found instability in the final decision process caused by weak confidence calibration.
- In **SEVADE**, studied how a self-evolving module updates agent teams through decoupled reasoning and evaluation, improving Macro-F1 by **6.29%** while revealing the need for adaptive retrieval in complex contexts.
- Developed **RAM-SD** (Retrieval-Augmented Multi-Agent Framework) with a “Learn-Before-Reason” strategy — pre-retrieving semantically similar training cases and applying contextual augmentation to guide reasoning, yielding a **+7.01%** Macro F1 improvement over gpt-4o+COG baselines.
- Contributed to **A Structured Path to Empathetic Response Generation with Multi-Agent Models**, a five-stage empathy model (Trigger–Response–Appraisal–Cause–Emotion), and built an LLM-based evaluation framework for measuring emotional coherence and human-like empathy generation.

Deep Research Agent & Self-Evolving Agents May 2025 – May 2026 Remote
Yale University — [Gerstein Lab](#) (Advisor: [Dr. Mark Gerstein](#))

- Helped to develop a comprehensive **Agent Knowledge Base (AKB)** for cross-framework reasoning, consolidating thousands of agent trajectories with both human-curated and LLM-generated analyses.
- Led large-scale reproduction experiments validating AKB’s effectiveness across smolagent, SWE-Agent, OpenHands, and OWL under the GAIA, SWE-bench, and GPQA benchmarks.
- Confirmed AKB’s consistent performance gains (up to **+18.7pp** in pass@3 on GAIA) and identified critical steps for deep research task decomposition and cross-domain knowledge transfer.
- Constructed a **self-evolving coding agent framework** with a dedicated Proposer Agent that autonomously generates and evaluates code modifications, enabling continual refinement and faster evolution cycles.
- Designed a data-centric, self-evolving agent framework that generates task-specific data (MCP tasks) and fine-tunes open-source LLMs via reinforcement learning (GRPO).

Acoustic Scene Classification & Accent Recognition Sep 2024 – Jun 2025 Suzhou, China
XJTLU — Audio Intelligence Lab (Advisor: [Dr. Shengchen Li](#))

- Designed and implemented the AdapTF-SepNet framework for multi-device acoustic scene classification.
- Developed an AudioSet-driven adaptive pre-training strategy with a transfer module aligning spectral and temporal filters to device-specific distributions.
- Engineered an end-to-end PyTorch Lightning pipeline integrating MixStyle, MixUp, and knowledge distillation to enhance generalization under unseen device domains.
- Conducted extensive ablation experiments showing **+7.61%** robustness and accuracy gains across real-world device splits.
- Contributed to accent recognition research using HuBERT embeddings and contrastive learning, achieving a **15%** accuracy boost via feature augmentation and domain adaptation.