Lingrui Mei (梅灵睿)

Phone:(+86)18512377787
GitHub:github.com/Meirtz
Email:meilingrui22@mails.ucas.ac.cn
Scholar: Google Scholar (1500+ citations)
Website:me.meirtz.com/about
Location: Beijing

Education

Institute of Computing Technology, Chinese Academy of Sciences |*Master's and PhD Student / Research Focus: Large Language Models* 2022.09—2027.07
State Key Laboratory of AI Safety | Computer Science and Technology

Chongqing Jiaotong University |*Bachelor of Engineering* 2015.09—2020.06
School of Civil Engineering | Civil Engineering and Computer Science & Technology

Work Experience

Tencent |*Large Language Model Department / T3 / Project Up (青云计划) Research Intern* 2025.11—Present

- **Built the verifiable RLVR data flywheel and sandbox environments for Computer Using Agents**: proposed an agentic data-evolution method to construct scalable agent tasks with matched asset sets, and closed the loop of scalable environment generation and maintenance, action execution, trajectory replay, automatic acceptance, and failure filtering, yielding a high-quality training loop for RFT/RL and participating in training.
- **Participated in merged OPD production training of Tencent Hunyuan Hy 2.0/3.0 and the subsequent fast-thinking model**: responsible for data mixing ratios, training configurations, evaluation regressions, and anomalous-sample attribution; iterated training recipes based on loss/reward curves, capability evaluations, and case replay; explored experimental model-merge schemes.
- **Participated in humanities reward-model training and online A/B optimization experiments**: built preference data, user-feedback reward-model training, and calibrated evaluation pipelines; tracked Yuanbao online feedback and in-house subjective evaluations on humanities and social sciences, long-text understanding, and complex QA scenarios, feeding results back into data iteration to continuously improve answer quality in humanities domains.

SkyworkAI |*2050 Research / Reinforcement Learning Research Intern* 2025.05—2025.11

- **Worked on the data flywheel and SFT/RFT/RL training of Skywork Deep Research Agent v2**: built the deep-research task data flywheel, constructing hard tasks by diversity, uniqueness, verifiability, and difficulty; combined GRPO, dynamic curriculum learning, and clue-based dense rewards to improve deep information retrieval and complex task execution.
- **Worked on RFT/SFT training and inference-paradigm design and implementation**: designed data recipes, tool-call trajectories, and verifiable answer-extraction pipelines, optimizing the loop from teacher rollout and training-sample filtering to offline evaluation.
- **Co-designed parallel thinking and generative result-verification mechanisms**: extended test-time scaling via multi-candidate paths, tournament ranking, and entropy-adaptive pruning; achieved 38.7% on BrowseComp and 79.1 on GAIA, setting new SOTA on multiple agent benchmarks.

Apple |*GC AIML / NLP Research Intern* 2025.02—2025.05

- **Designed and trained domain-specific agent foundation models**: constructed training data and offline evaluation sets around domain tasks, optimizing instruction formats, tool-calling strategies, and reasoning pipelines to adapt models to vertical scenarios.

Microsoft Research Asia |*DKI Group / Microsoft AI / Research Intern* 2024.10—2025.02

- **Co-proposed AdaptFlow/NL-MAML, a workflow meta-optimization framework**: represents agentic workflows as learnable natural-language programs, constructs subtask-level inner-loop adaptation via task clustering, and learns generalizable workflow initializations through outer-loop reflection and aggregation; averaged 68.5 across 8 QA, code, and math-reasoning benchmarks, outperforming workflow-search baselines such as AFlow and ADAS.
- **Co-proposed VEM, an environment-free RL framework for GUI agents**: models GUI decision-making as offline state-action value estimation, first training a Value Environment Model with semantic-progress signals, then freezing the VEM to guide action-policy exploration with PPO, bypassing online interaction, next-state simulation, and handwritten rewards; achieved SOTA/near-online-RL performance on AITW at lower interaction cost.

Xiaohongshu (rednote) |*Business Technology Department / Recommendation Algorithm Engineer* 2024.09—2024.11

- **Optimized Xiaohongshu homepage recommendation algorithms with large language model representations**, including internal- and external-flow generative recommendation models.
- **Achieved a 1.5‰ CTR improvement on business metrics**, significantly improving user experience and platform effectiveness.

China Automotive Engineering Research Institute |*State Key Laboratory of Vehicle NVH and Safety Technology | Autonomous Driving Algorithm Engineer* 2020.05—2021.10

- **Designed and developed cabin signal perception and road information perception algorithms**, including sensor fusion, vehicle testing software development, model deployment, and hardware performance optimization.
- **Participated in formulating C-IASI (China Insurance Automotive Safety Index autonomous driving assessment standards)**, with related research empowering the National Intelligent Vehicle Integration System Testing Zone.

Tsinghua University |*Center for Brain-Inspired Computing / Beijing Innovation Center for Future Chips | Research Intern* 2019.07—2020.02, 2024.04—2024.07

- **Optimized convolutional spiking neural networks and developed an EEG processing pipeline**: rewrote operators and integrated them into Kaldi (+26.4%), outperforming LSTM on some detection tasks (+4.5%); supported brain-signal decoding and multimodal generation.

Research Interests

- **Agentic RL and verifiable post-training**: RLVR/RFT data construction, scalable environments, programmatic rewards, and training iteration loops for Computer Using Agents, Code Agents, and Deep Research Agents.
- **Long-context and memory mechanisms for LLMs**: differentiable working memory, conditional computation, and context engineering to improve reliability in long-horizon reasoning and complex tasks.
- **LLM safety and trustworthy alignment**: safe generation, hallucination/jailbreak detection, context faithfulness, and multi-modal RAG bias.

Research Experience

Large Language Models, Trustworthy AI, Language Model Agents (2023.11—Present)

First-Author Publications

[1] Gated Differentiable Working Memory for Long-Context Language Modeling. [**ACL 2026, Oral**]

[2] a1: Steep Test-time Scaling via Environment Augmented Generation. [**ACL 2026**]

[3] HiddenGuard: Fine-Grained Safe Generation with Specialized Representation Router. [**ACL 2026**]

[4] A Survey of Context Engineering for Large Language Models. [**HuggingFace Monthly Ranking #2, covered by MIT Technology Review and other media**]

[5] "Not Aligned" is Not "Malicious": Being Careful about Hallucinations of Large Language Models' Jailbreak. [**COLING 2025**]

[6] SLANG: New Concept Comprehension of Large Language Models. [**EMNLP 2024**]

[7] Separations and Allocation for Conditional Memory and Computation. [NeurIPS 2026, Pending]

[8] Can We Trust Learned Speculative-Decoding Controls? [EMNLP 2026, Pending]

Co-Authored Publications

[1] Reward and Guidance through Rubrics: Promoting Exploration to Improve Multi-Domain Reasoning. [**ICML 2026, Spotlight**]

[2] YuE: Scaling Open Foundation Models for Long-Form Music Generation. [**ICLR 2026**]

[3] Parameters vs. Context: Fine-Grained Control of Knowledge Reliance in Language Models. [**ICLR 2026**]

[4] AdaptFlow: Adaptive Workflow Optimization via Meta-Learning. [**EMNLP 2025**]

[5] Who is in the Spotlight: The Hidden Bias Undermining Multimodal Retrieval-Augmented Generation. [**EMNLP 2025, Oral**]

[6] Decoding by Contrasting Knowledge: Enhancing LLMs' Confidence on Edited Facts. [**ACL 2025**]

[7] Context-DPO: Aligning Language Models for Context-Faithfulness. [**ACL 2025**]

[8] EvoStealer: Differential Evolution for Prompt Template Stealing Against Text-to-Image Synthesis. [**ACL 2025**]

[9] Is Factuality Enhancement a Free Lunch For LLMs? Better Factuality Can Lead to Worse Context-Faithfulness. [**ICLR 2025**]

[10] Adaptive Token Biaser: Knowledge Editing via Biasing Key Entities. [**EMNLP 2024**]

[11] LPNL: Scalable Link Prediction with Large Language Models. [**ACL 2024**]

[12] Detecting Out-of-distribution Samples via Variational Auto-encoder with Reliable Uncertainty Estimation. [**Neural Networks 145, 199-208**]

Technical Skills

- **LLM Training and Post-Training**: SFT, RFT, RLVR, reward modeling, and merged OPD production training; proficient with Megatron-LM, MBridge, veRL, PyTorch, and Transformers.
- **Agent Data Flywheel**: task generation, environment-state construction, programmatic rewards/verifiers, teacher rollout, data filtering, benchmark evaluation, and training feedback loops.

- **Research Areas**: Agentic RL, Computer Using Agents, Code Agents, Deep Research Agents, long-context memory mechanisms, and LLM safety and trustworthy alignment.

Awards and Achievements

- **Academic Service**: Reviewer for top-tier conferences including NeurIPS 2025, ACL 2025, NAACL 2025, ICML 2024; ICLR 2025 Notable Reviewer.
- **Competition Awards**: University of Chinese Academy of Sciences Outstanding Student 2024-2025; UCAS Academic Scholarship; 2020 SEC China Circuit Rally Championship (Runner-up), 2018 International Collegiate Brain-Inspired Computing Competition (Third Prize, only undergraduate), HACKxTHU 2018 First Prize, HACKxFDU 2017 (2/32), HACKxSJTU 2017 (Second Place), RoboCup 2013 (Individual and Team World Champion), etc.
- **Other Titles**: China Automobile and Motorcycle Sports Federation Class A Circuit Racing Driver / Beijing China Automobile Federation Racing Club Endurance Driver / University of Chinese Academy of Sciences Special Photographer.