梅霊睿（Lingrui Mei）

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

学歴

中国科学院計算技術研究所 | 修士／博士一貫課程 / 研究テーマ：大規模言語モデル 2022.09—2027.07
AI アルゴリズム安全国家重点実験室 | 計算機科学／技術

重慶交通大学 | 学士（工学） 2015.09—2020.06
土木工程学院 | 土木工学／計算機科学

インターンシップ経歴

Tencent | 大規模言語モデル部 / *T3* / 青雲計画リサーチインターン 2025.11—現在

- **Computer Using Agent 向け検証可能 RLVR データフライホイールとサンドボックス環境の構築**：エージェンティックなデータ進化手法によるスケーラブルなタスク・アセット構成、環境の生成・維持、アクション実行、軌跡リプレイ、自動検収と失敗フィルタリングのループを整備し、RFT/RL 用の高品質学習データを供給、学習にも参画。
- **Tencent Hy 2.0/3.0 と後続の高速思考モデル（OPD 統合版）の本番学習**：データ配合、学習設定、評価リグレッション、異常サンプルの帰因分析。loss/reward・能力評価・ケースリプレイに基づく学習方策の反復、実験的モデルマージの検討。
- **人文系能力の報酬モデル学習とオンライン A/B 最適化**：選好データ構築、ユーザーフィードバック RM の学習・校正評価。人文社会・長文理解・複雑 QA 場面のオンラインフィードバックを追跡しデータ反復に還元。

SkyworkAI |*2050* 研究院 / 強化学習リサーチインターン 2025.05—2025.11

- **Skywork Deep Research Agent v2 のデータフライホイールと SFT/RFT/RL 学習**：多様性・一意性・検証可能性・難度に基づく高難度タスク構成。GRPO・動的カリキュラム学習・手がかり型 Dense Reward により深い情報検索と複雑タスク遂行能力を向上。
- **RFT/SFT データレシピと推論パラダイムの設計**：ツール呼び出し軌跡・検証可能な解答抽出、teacher rollout からサンプル選別・オフライン評価までのループを最適化。
- **並列思考と生成的結果検証の設計**：多候補パス・トーナメントランキング・エントロピー適応枝刈りによる test-time scaling。BrowseComp 38.7%、GAIA 79.1、複数のエージェントタスクで当時 SOTA。

Apple |*GC AIML / NLP* リサーチインターン 2025.02—2025.05

- **ドメイン特化エージェント基盤モデルの設計・学習**：学習データとオフライン評価セットの構築、指示形式・ツール呼び出し戦略・推論チェーンの最適化による垂直領域適応。

Microsoft Research Asia (MSRA) |*DKI Group / Microsoft AI* | リサーチインターン 2024.10—2025.02

- **ワークフローメタ最適化フレームワーク AdaptFlow/NL-MAML を共同提案**：エージェンティックワークフローを学習可能な自然言語プログラムとして表現し、サブタスク単位の inner-loop 適応と outer-loop 反省・集約により汎化可能な初期化を学習。QA・コード・数学 8 ベンチマーク平均 68.5 で AFlow/ADAS 系を上回る。
- **環境非依存 GUI エージェント強化学習フレームワーク VEM を共同提案**：オフライン状態行動価値推定と凍結 VEM による PPO 誘導で、オンライン対話・次状態シミュレーション・手書き報酬を回避。AITW で低コストに SOTA 級性能。

小紅書（rednote） | ビジネス技術部 / 推薦アルゴリズムエンジニア 2024.09—2024.11

- **LLM 表現によるホームフィード生成的推薦の最適化**（内部・外部フィード）。ビジネス指標 CTR +1.5‰。

中国自動車工程研究院（CAERI） | 自動車 *NVH* ／安全技術国家重点実験室 / 自動運転アルゴリズムエンジニア 2020.05—2021.10

- **車室・道路知覚アルゴリズムの設計開発**：センサーフュージョン、車載試験ソフトウェア、モデルデプロイとハードウェア最適化。C-IASI 自動運転評価基準の策定に参画。

清華大学 | 類脳計算研究センター / 未来チップ技術高精尖研究センター / リサーチインターン 2019.07—2020.02, 2024.04—2024.07

- **畳み込みスパイキングニューラルネットワークの最適化と EEG パイプライン構築**：演算子を書き直し Kaldi に統合（+26.4%）、一部検出タスクで LSTM を上回る（+4.5%）。脳信号復号とマルチモーダル生成を支援。

研究興味

- **エージェンティック RL と検証可能ポストトレーニング**：Computer Using Agent・Code Agent・Deep Research Agent 向けの RLVR/RFT データ構成、スケーラブルな環境、プログラマティック報酬、学習反復ループ。
- **LLM の長文脈と記憶機構**：微分可能ワーキングメモリ、条件付き計算、長期推論の信頼性を高めるコンテキストエンジニアリング。

- **LLM の安全性と信頼できるアラインメント**：安全生成、幻覚／ジェイルブレイク判別、文脈忠実性、マルチモーダル RAG バイアス。

研究業績

大規模言語モデル・信頼できる *AI*・言語モデルエージェント（*2023.11*—現在）

筆頭著者

[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 月間ランキング 2 位、MIT Tech Review China などが報道**]

[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, 査読中]

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

共著

[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**]

スキル

- **LLM 学習・ポストトレーニング**：SFT、RFT、RLVR、報酬モデル、統合版本番学習。Megatron-LM、MBridge、veRL、PyTorch、Transformers。
- **エージェントデータフライホイール**：タスク生成、環境状態構成、プログラマティック報酬／検証器、teacher rollout、データ選別、ベンチマーク評価と学習のフィードバックループ。
- **研究領域**：エージェンティック RL、Computer Using Agent、Code Agent、Deep Research Agent、長文脈記憶機構、LLM 安全・信頼アラインメント。

受賞と学術活動

- **学術活動**：NeurIPS 2025、ACL 2025、NAACL 2025、ICML 2024 など主要会議の査読委員。ICLR 2025 Notable Reviewer。
- **受賞**：中国科学院大学優秀学生（2024–2025）、同学業奨学金。2020 SEC 中国サーキットラリー選手権準優勝、2018 国際大学生類脳計算コンテスト三等賞（唯一の学部生）、HACKxTHU 2018 一等賞、HACKxFDU 2017（2/32）、HACKxSJTU 2017 二位、RoboCup 2013 個人・団体世界チャンピオン。
- **その他**：中国自動車モータースポーツ連合会 A 級サーキットドライバー／北京 CAMF レーシングクラブ耐久ドライバー／中国科学院大学特約フォトグラファー。