Sijie Li

sijieli@andrew.cmu.edu GitHub Google Scholar

Education

Peking University Sep. 2022 – Jun. 2026
B.S. in Information and Computing Science, Turing Class
GPA: 3.806/4.000 **(Top 8%)** · Undergraduate advisor: Prof. Di He

Research Profile

Research focuses on **LLM agents operating over long-horizon machine-learning development workflows**. The work combines process-level evaluation of human–agent planning, community-driven agent systems, budget-aware experimental design for scaling laws, and iterative reasoning architectures to study how agents explore, validate, and allocate effort. Publications include co-first-authored work at ICLR 2026 and COLM 2026, with oral presentations at EACL 2026 and Sci-FM @ COLM 2026.

Research Experience

Carnegie Mellon University Feb. 2025 – Present
Student Research Intern, advised by Prof. Yiming Yang
· Conduct research on autonomous ML engineering, human–agent planning, and efficient experimental design for foundation models.
· Lead or co-lead CoMind and *Spend Less, Fit Better*, and contribute to TraceML, translating research ideas into evaluation frameworks, algorithms, and open-source artifacts.

Publications

Spend Less, Fit Better: Budget-Efficient Scaling Law Fitting via Active Experiment Selection
Sijie Li*, Shanda Li*, Haowei Lin, Weiwei Sun, Ameet Talwalkar, Yiming Yang
COLM 2026 · Sci-FM @ COLM 2026 Oral
· Formulated scaling-law fitting as budget-aware sequential experimental design, selecting among candidate runs with heterogeneous costs to improve extrapolation in a high-cost target region.
· Designed and implemented an uncertainty-aware acquisition method evaluated on **8 tasks and 65 scaling-law instances**; it outperformed classical baselines and approached full-set fitting performance using **roughly 10% of the full budget**.

CoMind: Towards Community-Driven Agents for Machine Learning Engineering
Sijie Li*, Weiwei Sun*, Shanda Li, Ameet Talwalkar, Yiming Yang
ICLR 2026
· Led the design of MLE-Live, a live evaluation framework spanning **75 historical and 8 ongoing Kaggle competitions** that simulates community-driven ML research through shared discussions and code.
· Proposed CoMind, a multi-agent system that combines collective-knowledge utilization with iterative parallel exploration; achieved a **36% medal rate** and **outperformed 92.6% of human competitors** on average, reaching the top 5% on three live leaderboards and the top 1% on one.

TraceML: An Empirical Analysis of Human-Agent Planning in Machine Learning Development
Jiarui Yan, Weiwei Sun, **Sijie Li**, Wenhan Li, Yiming Yang
NeurIPS 2026 Evaluations & Datasets Track (under review) · WAB @ COLM 2026 Poster
· Contributed to the design of the human-trajectory collection protocol and downstream analysis under a unified version-level schema covering actions, intent, edit size, timestamps, scores, and score effects.
· Analyzed **4,465 human trajectories across 134 competitions** and a seven-competition matched set of 430 human and 207 agent trajectories, identifying systematic differences in exploration, validation, ensembling, and effort allocation.
** Equal contribution.*

Publications (continued)

- **Enhancing Auto-regressive Chain-of-Thought through Loop-Aligned Reasoning**
  Qifan Yu, Zhenyu He, **Sijie Li**, Xun Zhou, Jun Zhang, Jingjing Xu, Di He
  EACL 2026 Oral
  · Designed controlled evaluations of Loop Transformers and loop-aligned autoregressive models on longest-increasing-subsequence, arithmetic, and edit-distance tasks.
  · Developed and evaluated CoT models aligned with noisy intermediate states of Loop Transformers, connecting latent iterative computation with explicit step-by-step reasoning.
- **GarmentLab: A Unified Simulation and Benchmark for Garment Manipulation**
  Haoran Lu*, Ruihai Wu*, Yitong Li*, **Sijie Li**, Ziyu Zhu, Chuanruo Ning, Yan Shen, Longzan Luo, Yuanpei Chen, Hao Dong
  NeurIPS 2024
  · Contributed to a multi-physics simulation engine integrating PBD and FEM methods across **20 manipulation tasks**, including garment–rigid, garment–fluid, and garment–human interactions.
  · Designed a teleoperation pipeline that captures human hand poses and retargets them to dexterous robotic hands in simulation.
  ** Equal contribution.*

Academic Service & Outreach

- **Turing Class Research Forum — Volunteer Organizer & Speaker** 2024 – 2026
  · Coordinated invited speakers, promotional materials, and event logistics for an annual cross-institutional forum involving Peking University, Wuhan University, and the National University of Singapore.
  · Delivered a research tutorial on *Spend Less, Fit Better*, introducing budget-aware scaling-law experiments to students and faculty across participating Turing programs.
- **Competitive Programming Mentor, Sichuan, China** 2022 – 2025
  · Devoted more than four months across three undergraduate years to mentoring **100+ middle- and high-school students**; selected training problems, taught algorithms, and organized practice examinations.
  · Provided individualized performance analysis and study plans; dozens of students later earned first prizes in NOIP and CSP.
- **Academic Reviewer** 2025 – 2026
  · Reviewed **10+ submissions** for MTI-LLM @ NeurIPS 2025 and Sci-FM @ COLM 2026.
  · Assessed methodological soundness, experimental evidence, and presentation clarity, providing detailed and actionable feedback to authors.

Honors and Awards

- **Academic Excellence Scholarship** × **2** (Top 10%, Peking University) 2025
- **Excellence in Learning Scholarship** (Top 15%, Peking University) 2023
- **Gold Medal (8th in China)**, 47th ACM ICPC Asia Hefei Regional Contest 2022
- **Gold Medal (20th in China)**, China Collegiate Programming Contest, Weihai 2022
- **Silver Medal (64th in China)**, National Olympiad in Informatics (NOI) 2021
- **Gold Medal (5th in China)**, CCF Winter Camp 2021