Miri (Emma) Liu

Ph.D. Student | Siebel School of Computing and Data Science | University of Illinois, Urbana-Champaign

E-mail miri3@illinois.edu
Website emmatliu.github.io
LinkedIn emmatliu
Github emmatliu

PhD student working on AI for scientific discovery, including evaluating conceptual novelty in fine-grained, interpretable ways and building memory systems for long-horizon collaborative agents.

Education

University of Illinois, Urbana-Champaign Sep. 2025 - Present
Doctor of Philosophy | Computer Science

- Advisors: Dr. ChengXiang Zhai and Dr. Jiawei Han

University of California, Los Angeles Sep. 2021 - Jun. 2025
Bachelor of Science | Computer Science, Mathematics

Publications

Conference & Workshop Proceedings:

[1] Lu, L-C., **Liu, M.**, Lu, P-C., Tian, Y., & Peng, N. Rethinking creativity evaluation: A critical analysis of existing creativity evaluations. EACL 2026.

[2] Tian, Y.[*], Huang, T.[*], **Liu, M.**, Jiang, D., Spangher, A., Chen, M., May, J., & Peng, N. Are large language models capable of generating human-level narratives?. EMNLP 2024. **Outstanding Paper Award**.

[3] Acharya, N.[*], **Liu, E.**[*], Lucas, A.[*], & Lazar, A. Optimizing the Exa.TrkX inference pipeline for manycore CPUs. Connecting the Dots Workshop 2022.

[4] Lazar, A., et al. (including **Liu, E.**[*]). Accelerating the inference of the Exa.TrkX pipeline. ACAT 2021.

Preprints:

[5] **Liu, M.**, & Zhai, CX. (2026). An axiomatic benchmark for evaluation of scientific novelty metrics.

[*]*Equal contribution.*

Honors and Awards

Computational Science Graduate Fellowship, Department of Energy 2026-2030
4-year award; one of ∼30 awarded annually.

AI PhD Fellowship, Amazon 2025-2027
2-year award; one of 100 fellows selected.

Outstanding Paper Award, EMNLP 2024
"Are Large Language Models Capable of Generating Human-Level Narratives?"

Research Experience

Applied Scientist Intern May 2026 - Present
Amazon

PhD Student Researcher Sep. 2025 - Present
TIMAN Group | University of Illinois, Urbana-Champaign
Principal Investigator: Dr. ChengXiang Zhai

- **Scientific novelty:** Establishing robust evaluation methods for scientific novelty metrics and designing a new, interpretable metric to target existing weaknesses.

PhD Student Researcher Sep. 2025 - Present
Data Mining Group | University of Illinois, Urbana-Champaign
Principal Investigator: Dr. Jiawei Han

- **Agent memory:** Identifying failure modes in long-horizon agent state tracking that are not captured by retrieval quality, and designing a state-unit memory system.

Undergraduate Student Researcher Oct. 2023 - Jun. 2025
PLUS Lab | University of California, Los Angeles
Principal Investigator: Dr. Nanyun Peng

- **Narrative generation:** Devised methods to compare LLM and human narrative outputs, crafted prompts and the pipeline for LLM generations, and curated a dataset through Wikipedia scraping.
- **Creativity evaluation:** Contributed to a critical evaluation of automatic creativity metrics, revealing that current measures capture only surface-level novelty rather than true conceptual creativity.

Undergraduate Student Researcher Sep. 2021 - May 2022
XSEDE EMPOWER | Youngstown State University
Principal Investigator: Dr. Alina Lazar

- Applied high-performance computing techniques to optimize the Exa.TrkX inference pipeline on manycore CPU architectures, improving parallel efficiency.

Teaching Experience

University of California, Los Angeles

- Grader (CS 181, ECE M148) Jun. 2024 - Aug. 2024
  Graded assignments and provided weekly feedback to classes of 75 and 115 students, respectively.
- Learning Assistant, Software Construction (CS 35L) Mar. 2022 - Mar. 2024
  Held office hours and gave weekly instruction to a section of 50 students.

Industry Experience

Data Platforms Intern (SWE) Jun. 2023 - Sep. 2024
Balyasny Asset Management | Chicago, IL

- Designed and deployed a metadata ingestion framework integrating internal systems, saving an estimated *$250K in developer hours* and processing *10,000+ data assets*.
- Implemented a robust proxy service to facilitate a traffic increase from *3 million to 25 million* daily requests.

EXPLORE Software Intern (SWE) Jun. 2022 - Sep. 2022
Microsoft | Redmond, WA

- Built a data pipeline expanding monitoring scope from *300 to 3 million* server units.

Academic and Volunteer Service

Sponsorships Committee | LA Hacks Sep. 2022 - Jun. 2024

- Helped secure $180,000 in sponsorships through outreach to hundreds of potential partners.
- Organized panels and mentor-recruitment efforts to support over 1,000 hackathon participants annually.

Skills

Programming Languages: Python, C++; familiar with Java and C#.

Machine Learning: PyTorch, Hugging Face Transformers; LLM prompting and generation pipelines; model and metric evaluation.

Data & Analysis: NumPy, Pandas, SciPy; large-scale data-pipeline development, metadata ingestion, and web scraping; workflow automation with Apache Airflow.

Research Methods: Evaluation and benchmark design, dataset curation, human-annotation studies.

Last updated in June 2026