Contact: siddarth.mamidanna@gmail.com

LinkedIn
X/Twitter
GitHub
Google Scholar

Siddarth Mamidanna

Education

University of California, Santa Cruz — Santa Cruz, CA
B.S. Computer Science; Minor in Applied Mathematics & Computer Engineering (Junior) — *Sep 2022 – Present*

- Advisors: Prof. Leilani H. Gilpin (AIEA Lab), Dr. Yilun Zhou
- Research focus: interpretability and safety of large language models

Research

LLM Interpretability & Safety

- Mechanistic interpretability of transformers; studying where and how computation is localized/routed and connecting circuits to reliability/safety.
- Co-first authored a large empirical study on LLM self-explanations (100+ citations).
- Recent work with Y. Zhou and Z. Yao: mental-math computation concentrates at the last token after brief cross-token transfer (*EMNLP 2025*).
- Broader LLM applications in education (USC ICT): fine-tuning vs. few-shot for automated short-answer grading (*AIED Workshop 2025*).

Publications

Reverse chronological order. *Equal contribution.
C: conference. **W**: workshop. **P**: pre-print. Highlighted work.

C1 **Siddarth Mamidanna**, Daking Rai, Ziyu Yao, Yilun Zhou. *All for One: LLMs Solve Mental Math at the Last Token With Information Transferred From Other Tokens.* Empirical Methods in Natural Language Processing (EMNLP), 2025. PDF.

W1 Joel Walsh*, **Siddarth Mamidanna***, Benjamin Nye, Mark Core, Daniel Auerbach. *A Comparison of Fine-Tuning and Few-Shot Approaches for AI-based Short Answer Grading.* AIED Workshop, 2025. PDF.

P1 Shiyuan Huang*, **Siddarth Mamidanna***, Shreedhar Jangam, Yilun Zhou, Leilani H. Gilpin. *Can LLMs Explain Themselves? A Study of LLM-Generated Self-Explanations.* arXiv, 2023. PDF.

Work Experience

HeyMarin.ai — Santa Cruz, CA
Founder, CEO — *Aug 2025 – Present*

- Built an email-native AI assistant (currently in private beta)
- Engineered backend infra using AWS; designed data pipelines, wrote agent logic, and wrote mcps/integrations
- Owned product roadmap, infra, and early user pilots.

AI Explainability & Accountability (AIEA) Lab, UCSC — Santa Cruz, CA
Undergraduate Researcher — *May 2023 – Present*

- First authored EMNLP 2025 paper in mechanistic interpretability, investigating LLM reasoning at a token level with Yilun Zhou and Ziyu Yao
- Co-first authored study on LLM self-explanations; led experiment design, metrics (comprehensiveness, sufficiency), and analysis.

MCHN Ventures — Palo Alto, CA
Hacker-in-Residence — *Jun 2025 – Aug 2025*

- Venture building at leading global automotive OEM
- Prototyped multi-agent LLM systems for mobility applications; emphasized reliability, latency, and safe model orchestration.
- Built evaluation harnesses and model-serving pipelines for real-time inference.

USC Institute for Creative Technologies Los Angeles, CA
Research Intern *Jun 2024 – Aug 2024*

- Worked with Prof. Ben Nye and Joel Walsh
- Constructed cleaned/augmented datasets from OpenTutor; implemented evaluation pipeline for automated grading.
- Compared QLoRA fine-tuning (Llama-3) vs. few-shot with API models for short-answer scoring
- Culminated in workshop paper accepted @ AIED 2025

Bay AI (Nonprofit) Cupertino, CA
Co-President & ML Instructor *May 2019 – Jun 2022*

- Delivered weekly 60–90 min AI lectures; authored hands-on exercises and demos for high-school audiences.
- Coordinated sponsors, volunteers, and events across the Bay Area.