



Capital One Fellows

**AY 2026-2027
Fellow Bios**

Columbia University



Jinjun Peng

Google Scholar

Project Title: Secure CodeLLM with Reasoning

Faculty Advisor: Dr. Baishakhi Ray

Bio: Jinjun Peng is a Ph.D. student in Computer Science at Columbia University, advised by Professors Baishakhi Ray and Junfeng Yang. His research has centered on code language models, including SWE-Spot, which builds small repository-expert models through repository-centric learning; CWEval, an outcome-driven benchmark evaluating both the functionality and security of LLM-generated code; and SemCoder, which trains code language models with comprehensive semantic reasoning. His current research interest is continual learning, i.e., how AI agents can effectively and efficiently acquire new knowledge and adapt to evolving environments.



Tao Long

[Google Scholar](#)

Project Title: Vibe-Working: Agentic Infrastructure for Humans to Become AI Supervisors for General Knowledge Work

Faculty Advisor: Dr. Lydia Chilton

Bio: Tao is a Computer Science PhD student at Columbia University, advised by Professor Lydia Chilton. Working on human–AI interaction, Tao’s research explores how humans collaborate with generative AI systems and AI agents over time, focusing on making AI tools more usable, useful, trustworthy, reliable and seamlessly integrated into everyday productivity practices. Specifically, Tao builds human–AI and agentic systems that reduce cognitive and temporal effort for challenging or complex tasks, offload work to AI while maintaining human ownership and authenticity and fit naturally into users’ existing processes for writers, developers, designers, event organizers and many other communities. Before starting their PhD, Tao earned a BS Summa Cum Laude from Cornell University.



Ved Sriraman

Google Scholar

Project Title: Understanding and Stabilizing Reinforcement Learning Dynamics in Post-Training Language Models

Faculty Advisor: Dr. Adam Block

Bio: Ved is a Computer Science PhD student at Columbia University, advised by Professor Adam Block. His research focuses on improving the capabilities and reliability of LLMs through post-training and test-time algorithms. Combining theoretical analysis with empirical investigation, Ved develops algorithmic interventions that come with provable guarantees while remaining relevant to modern language-model training. More broadly, he is interested in uncovering fundamental mathematical principles of learning in real-world environments and distilling them into methods that enable language models to reason. Before starting his PhD, Ved earned a BS in Computer Science from Cornell University.



University of Illinois



Jeonghwan Kim

Google Scholar

Project Title: Enhancing Real-World Grounding through Textual-Visual Representation Interleaving in Multimodal Large Language Models

Faculty Advisor: Dr. Heng Ji

Bio: Jeonghwan Kim is a Ph.D. candidate in Computer Science at the University of Illinois Urbana-Champaign, advised by Professor Heng Ji. His research focuses on multimodal foundation models that bridge fine-grained visual perception, reasoning, and embodied intelligence. Prior to his work on multimodal AI, he conducted research in natural language processing, including multi-hop question answering, retrieval-augmented generation, and numerical reasoning. His work has appeared in leading venues such as NeurIPS, ICLR, ACL, EMNLP, NAACL, and CVPR, including a NeurIPS 2025 Spotlight paper on part-level visual understanding in large multimodal models.



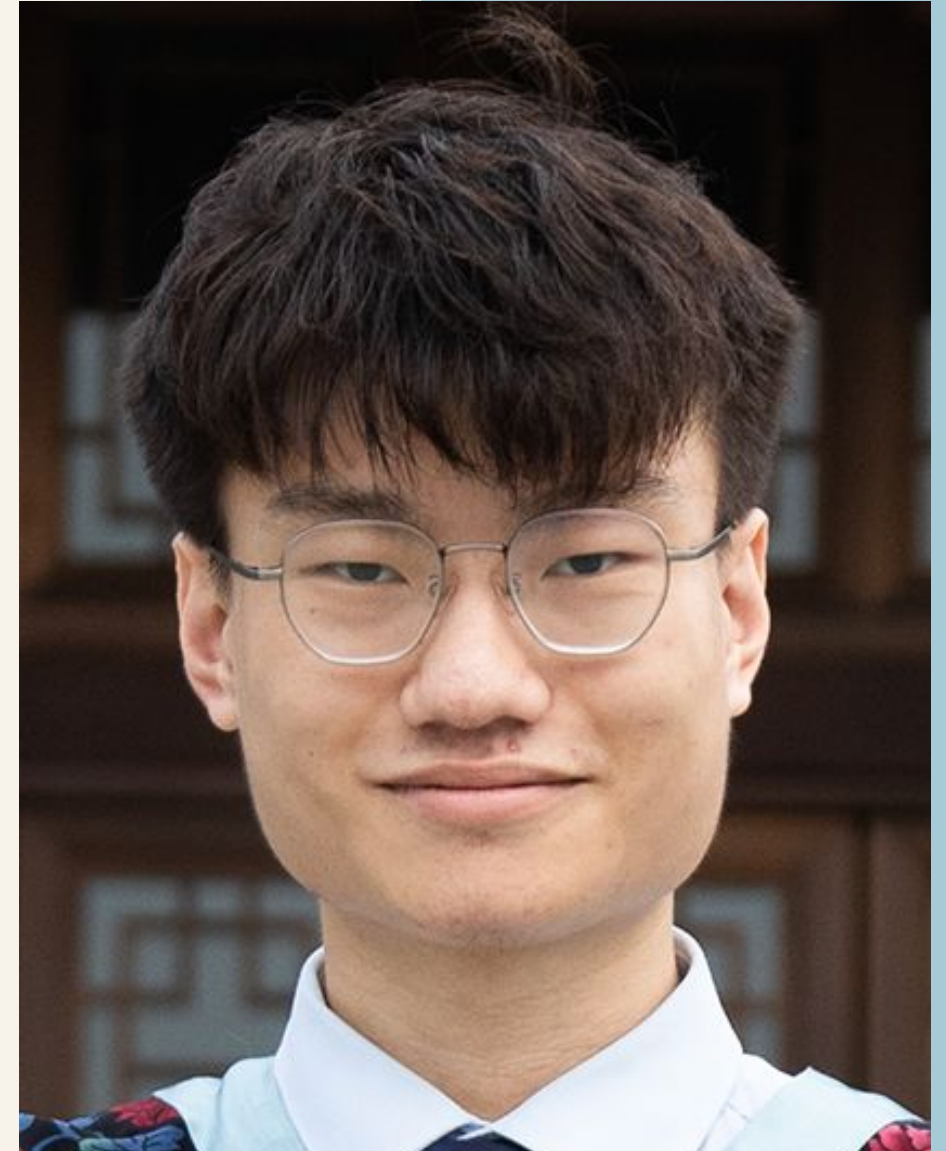
Yizhuo Chen

Google Scholar

Project Title: Utility-Preserving Data Sanitization for Generative AI Model Development

Faculty Advisor: Dr. Tarek Abdelzaher

Bio: Yizhuo Chen is a PhD candidate in Computer Science at the University of Illinois Urbana-Champaign, advised by Prof. Tarek Abdelzaher. His research centers on trustworthy machine learning and foundation models for sensing, with a particular focus on privacy-preserving data analytics and the safe use of sensitive data in generative AI. His work has appeared at venues including ICML and CVPR, and he has collaborated with industry research labs at JPMorgan Chase and Amazon. He received his B.Sc. from Zhejiang University in 2021.



Hyeonjeong (Amber) Ha

Google Scholar

Project Title: Toward Creative and Reliable Large Language Models

Faculty Advisor: Dr. Heng Ji

Bio: Hyeonjeong is a third-year PhD student at UIUC, advised by Professor Heng Ji. Her research enhances the visual perception of multimodal LLMs through fine-grained, structured understanding, aiming to bridge human-like perception and trustworthy reasoning for more reliable real-world applications. She has published as first author in top-tier conferences including ACL and NeurIPS, and actively contributes to the research community as a reviewer for leading venues.



University of Virginia



Mia Yuan

Google Scholar

Project Title: Integrative Decoding for Reliable Large Language Model Reasoning in Financial Services

Faculty Advisor: Dr. Sheng Li

Bio: Hanzhang “Mia” Yuan is a fourth-year PhD student in Data Science at UVA, advised by Dr. Sheng Li. Her research focuses on improving the reasoning and reliability of large language and multimodal models. As a Capital One PhD Fellow, she is extending this work to financial tabular reasoning, with an emphasis on helping models interpret structured information and perform complex numerical reasoning more faithfully. Her broader goal is to develop trustworthy AI systems that can support reliable decision-making in real-world financial applications.



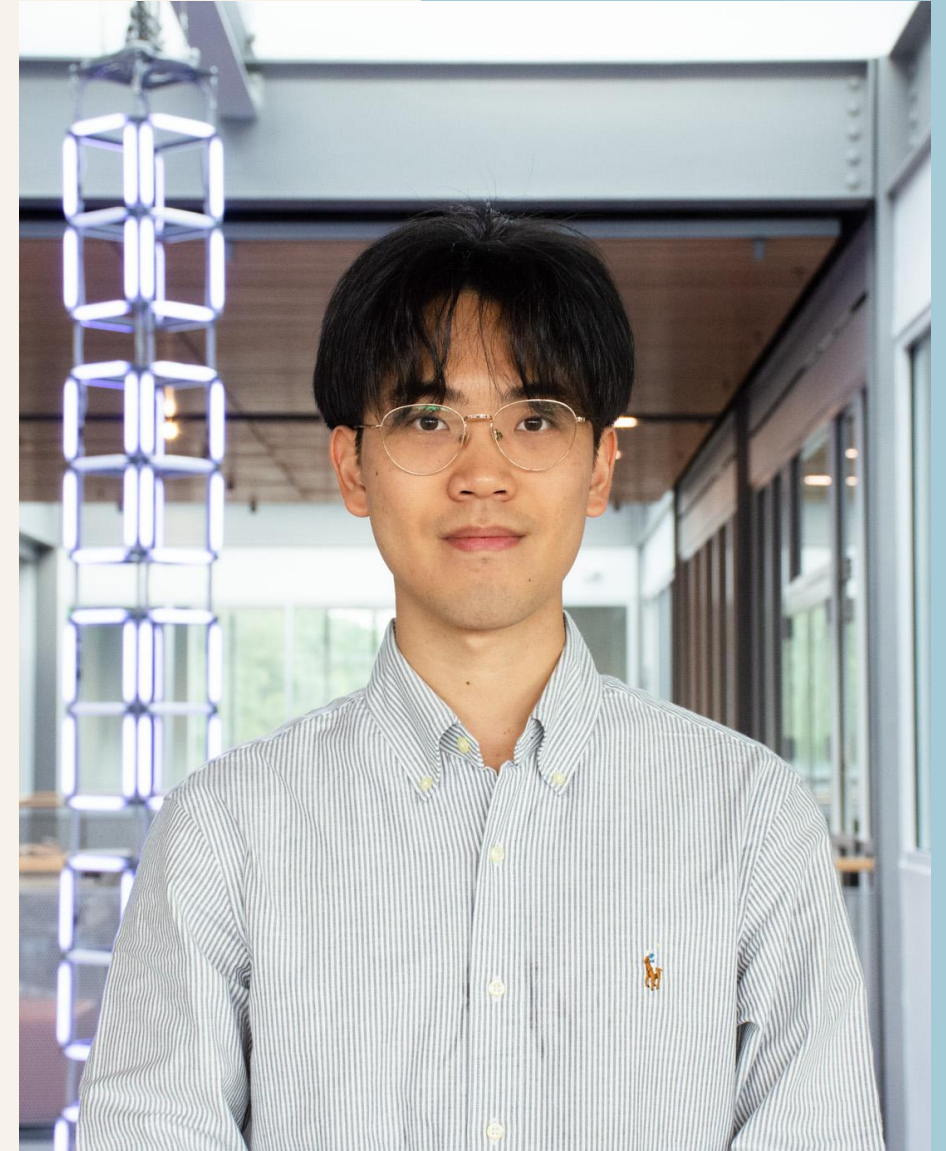
Ding Zhang

Google Scholar

Project Title: Towards Multimodal Graph-Language Reasoning

Faculty Advisor: Dr. Chirag Agarwal

Bio: Ding is a second-year PhD data science student at UVA advised by Dr. Chirag Agarwal. He is interested in modelling such graph data using graph representation learning techniques: e.g., applying GNNmodels. Additionally, as AI systems become increasingly integrated into critical decision-making processes, such as health care, finance and public policy, it is essential that these systems are not only accurate but also interpretable and reliable. Building models that can clearly explain their predictions is crucial for fostering user trust and ensuring ethical deployment. This area is the key to bridging the gap between technical innovation and real-world, responsible adoption of AI.



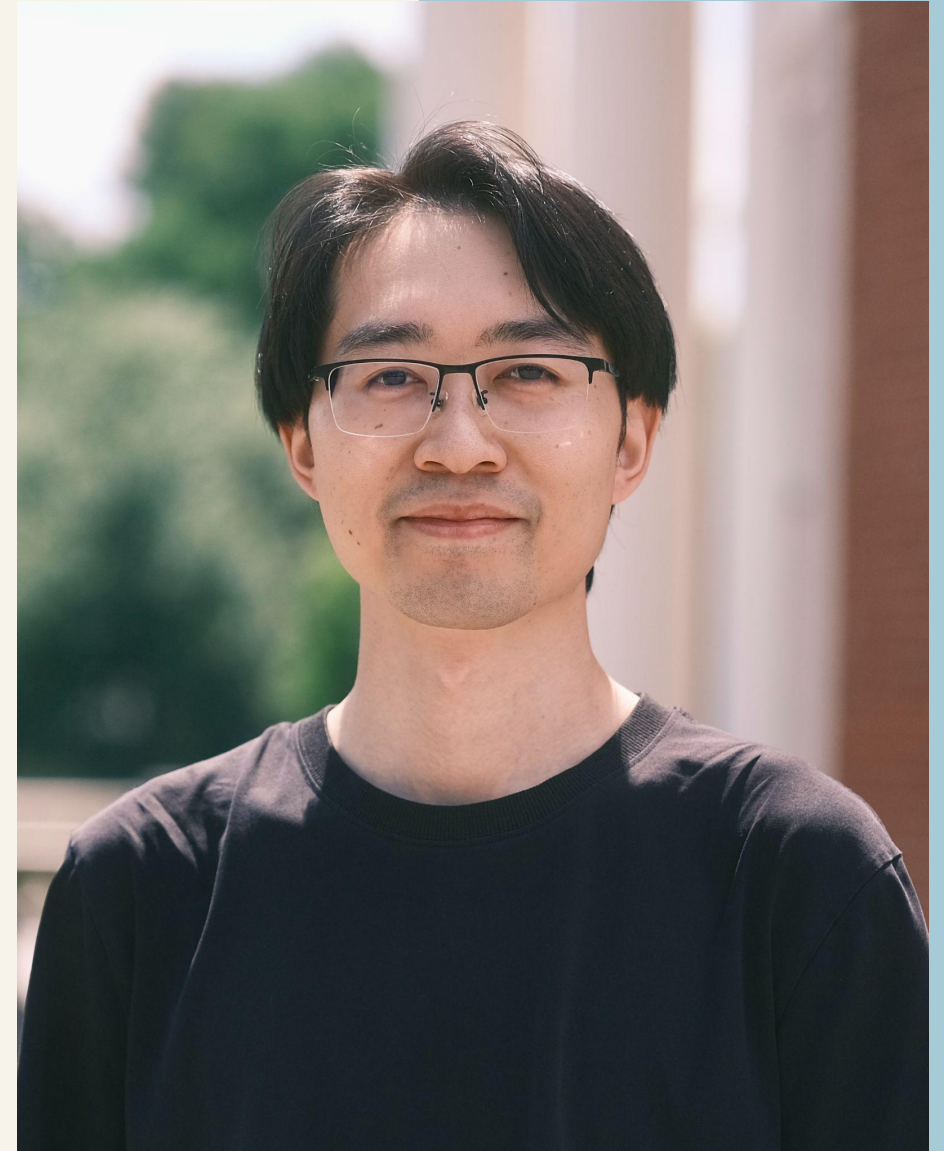
Zhepei Wei

Google Scholar

Project Title: Toward Trustworthy and Efficient Agentic AI

Faculty Advisor: Dr. Yu Meng

Bio: Zhepei Wei is a Ph.D. candidate in the Computer Science department at the University of Virginia, advised by Professor Yu Meng. Zhepei has held research positions at Microsoft Research, Meta and Amazon, working on large language models (LLMs). His first-authored research papers have been published in top-tier venues in the fields of machine learning and artificial intelligence (e.g., NeurIPS, ICML, ICLR) with over 1,400 citations on Google Scholar. He is also a recipient of the UVA Copenhaver Charitable Trust Bicentennial Fellowship and the John A. Stankovic Outstanding Graduate Research Award.



Zhenyu Lei

Google Scholar

Project Title: Towards Efficient, Reliable, and Multimodal AI Reasoning in Financial Services

Faculty Advisor: Dr. Jundong Li

Bio: Zhenyu Lei is a Ph.D. student in electrical and computer engineering at the University of Virginia, advised by Professor Jundong Li. Before joining UVA, he earned his B.S. in physics (Honors Graduate) from Xi'an Jiaotong University. His research centers on making LLM reasoning more efficient and reliable. Specifically, he works on reasoning distillation, which compresses powerful models into smaller ones that retain strong reasoning ability, and on reasoning editing, which corrects failures in model behavior and enables LLMs to reason reliably over their lifetime. Building on this foundation, he is excited to extend his work to the financial domain, where efficient and trustworthy reasoning is critical for real-world decision-making. He has published over twenty papers at leading venues including ICLR, AAAI, and ACL, with three oral presentations.



University of Southern California



Kien Nguyen

Google Scholar

Project Title: Project Title: Dynamic Schema-guided Monte Carlo for Text-to-SQL Agents on Massive Databases

Faculty Advisor: Dr. Paul Bogdan

Bio: Kien is a fourth-year PhD student in Computer Engineering at USC, advised by Professor Paul Bogdan. His prior research focused on machine learning and graph representation learning. He is currently interested in leveraging graph learning techniques for LLM agent research, including self-evolving and multi-agent systems. In collaboration with Capital One, he is developing advanced agent systems for data analytics and decision-making tasks across businesses. He has published his work in and served as a reviewer for prestigious venues such as AAAI, WWW, and KDD.



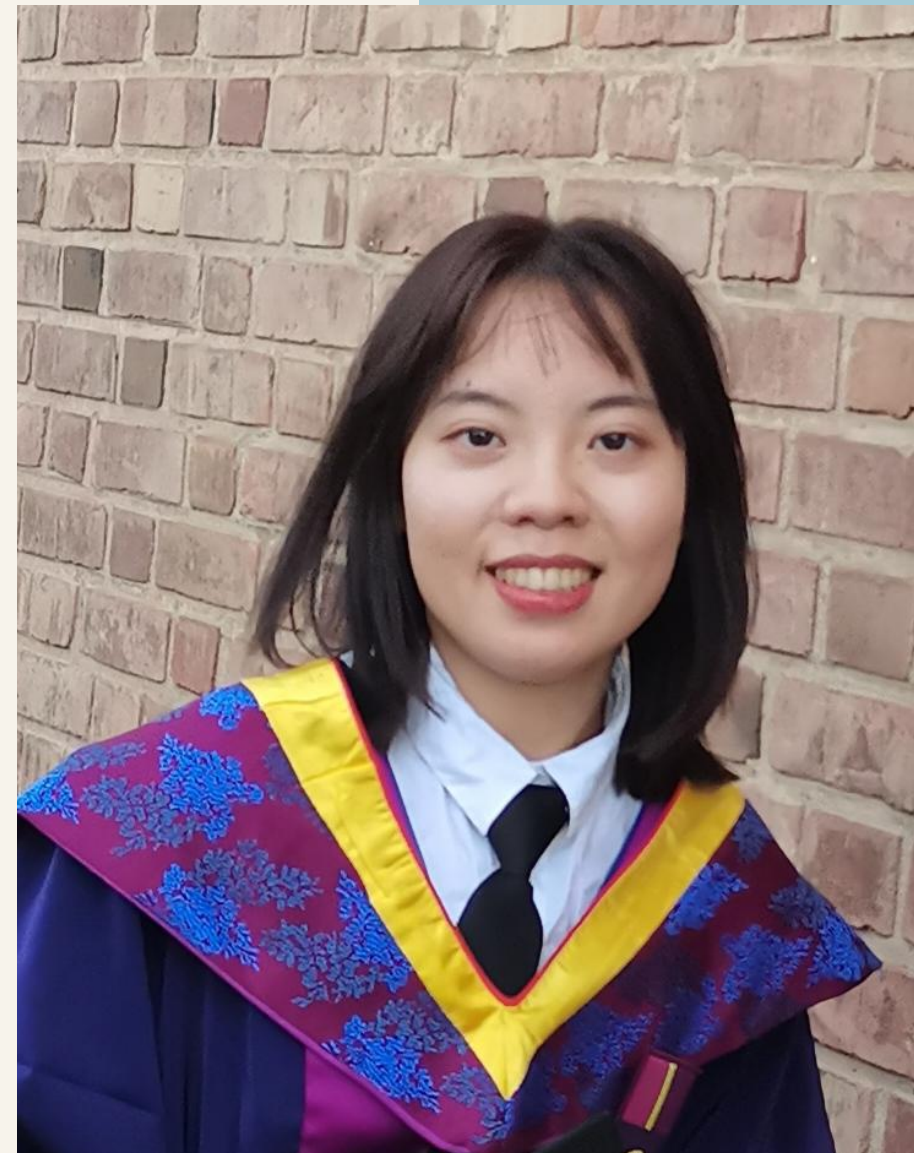
Yuxin Yang

Google Scholar

Project Title: Query-Adaptive Graph Traversal and Persistent Investigation Memory for Financial Knowledge Extraction

Faculty Advisor: Dr. Viktor Prasanna

Bio: Yuxin Yang is a fourth-year PhD student in Computer Science at the University of Southern California, advised by Professor Viktor Prasanna. Her research focuses on building reliable and efficient LLM agents, with particular emphasis on controlled test-time scaling—enabling agents to adaptively decide how to search, reason, verify and manage context during long-horizon tasks. In collaboration with Capital One, she developed SPARC-RAG, an agentic retrieval-augmented generation framework that dynamically allocates computation across sequential and parallel search while consolidating evidence in a unified context state. Her broader research spans post-training, deep research and coding agents, as well as graph learning, with the goal of making AI systems more capable, efficient and trustworthy in complex decision-making settings. Before joining USC, she received her B.E. degree from Tsinghua University.



Md Abrar Jahin

Google Scholar

Project Title: FinSemKG: A Semantic Mediation Layer for Grounded, Auditable LLM Question Answering over Financial Tables

Faculty Advisor: Dr. Craig Knoblock

Bio: Abrar is a second-year PhD student in Computer Science at USC, co-advised by Professor Craig Knoblock and Professor Jay Pujara. He is also an AI researcher with the Center on Knowledge Graphs at USC's Information Sciences Institute. His research focuses on structured and trustworthy AI, particularly knowledge representation, semantic table understanding, geometric deep learning, and grounded LLM systems. As a Capital One Fellow, he is developing FinSemKG, a semantic mediation layer for grounded, provenance-aware, and auditable LLM question answering over heterogeneous financial tables. He also contributes to DARPA/USGS- and AFRL-funded projects involving scientific knowledge extraction, knowledge graph construction, and critical mineral discovery. Abrar has authored more than 20 peer-reviewed publications that have received over 700 citations. His distinctions include the Viterbi Graduate School Fellowship and two Highly Commended distinctions from The Global Undergraduate Awards, placing his undergraduate research in the top 10% globally. He also serves as a program committee member and reviewer for leading AI and machine learning venues, including AAAI, ICDM, NeurIPS, ICML, and TMLR, and received an ICML 2026 Gold Reviewer Award. Before joining USC, he earned his BSc in IPE from KUET and served as a visiting researcher at the Okinawa Institute of Science and Technology in Japan.



Shahab Sepehri

Google Scholar

Project Title: XFinAgent: Financial Analyst Agents with Sparse-Probe Control, Multi-Agent Explainability, and Robust Evaluation

Faculty Advisors: Dr. Mahdi Soltanolkotabi and Dr. Stephen Tu

Bio: Shahab is a fourth-year PhD candidate advised by Prof. Mahdi Soltanolkotabi at the University of Southern California. His research focuses on developing new algorithms for multimodal reasoning in vision-language and diffusion models, enabling them to better integrate, generate, and reason across modalities. He also works on time-series foundation models, developing more effective and efficient tokenization schemes for diverse temporal data.



Asal Mehradfar

[Google Scholar](#)

Project Title: Uncertainty Quantification for Multi-Agent LLM Systems: Decomposition and Failure Localization

Faculty Advisor: Dr. Salman Avestimehr

Bio: Asal Mehradfar is a fourth-year PhD candidate in Electrical Engineering at the University of Southern California, advised by Professor Salman Avestimehr. Her research focuses on developing AI and machine learning methods for engineering systems and scientific discovery. She is particularly interested in building reliable, domain-informed learning and generative frameworks that integrate data, scientific knowledge, and human expertise. Her work spans financial market modeling, electronic design automation, and biomedical systems.



Xinyue Cui

Google Scholar

Project Title: Multi-Signal Reward Modeling for Robust and Trustworthy AI

Faculty Advisor: Dr. Swabha Swayamdipta

Bio: Xinyue is a third-year PhD student at USC, advised by Professor Swabha Swayamdipta. Her research focuses on advancing AI safety and trustworthiness through a data-centric lens, investigating how the quality, integrity, and diversity of training data shape model reliability, robustness, and susceptibility to manipulation. Prior to her PhD, she received her MS in Computer Science from USC and her BS in Computational Mathematics from UCLA.



Zarif Ikram

Google Scholar

Project Title: Belief-Weaver: Probabilistic Belief Transformers for Synergizing Language, Knowledge, and Rich Feedback

Faculty Advisor: Dr. Paria Rashidinejad

Bio: Zarif Ikram is a 2nd-year PhD student in Computer Engineering at USC, advised by Paria Rashidinejad. His research interests lie at the intersection of world models and scalable oversight, with a major theme being the principled understanding of “why” in training foundation models. Zarif earned a BS from BUET in 2025 and interned at Microsoft Research in 2026. Beyond research, he is a classically trained vocalist and multi-instrumentalist.

More information about his work is available on his [personal website](#).



Amin Banayeeanzade

Google Scholar

Project Title: Unlocking Diverse Thinking and Planning in AI Agents

Faculty Advisor: Dr. Sai Praneeth Karimireddy

Bio: I'm a Ph.D. student in Computer Science at USC. I study the trustworthiness of AI systems with a particular focus on diversity in large language models (LLMs) and agentic systems. I did my Master's and undergrad at Sharif University of Technology.

