Ali Keramati

+1 (951) 258-0875 a.kera@uci.edu alikera.github.io alikera ali-keramati Google Scholar

Education

University of California Irvine **Jan. 2025 – Now**

Ph.D in **Education Data Science** (GPA: 4/4) *Irvine, CA, USA*

- **Relevant Courses:** *Neural Networks and Deep Learning, Deep Generative Models, Neural Architecture of Language, Language Models in Cognitive Science, Learning Analytics Practicum, Qualitative Research Methods*

University of Tehran **Sep. 2019 – Jul. 2024**

B.Sc. in **Computer Engineering** (GPA: 18.51/20) *Tehran, Iran*

- **Major Degree Courses:** *Discrete Mathematics, Data Structures and Algorithms, Artificial Intelligence, Machines and Language Theory, Engineering Probability and Statistics, Internet Engineering, Foundations of IT, Large Language Models*
- **Minor Degree in Educational Science Courses:** *Educational Psychology, Instructional Design, Curriculum Development, Intro to Psychological Tests, Teaching Methods, Qualitative and Quantitative Research Methods*

Research Interests

- Multi-Agent Systems
- LLM Tuning (PEFT/LoRA)
- Alignment (RLHF/RLAIF)
- LLM-as-Judge
- Human-AI Collaboration
- Autonomous Learning Agents

Work Experience

Research Scientist Intern at Harvard University **Jul. 2025 - Sep. 2025**

Learning Media Lab, PI: Prof. Y. Xu *Cambridge, MA*

- Built web-based math experiments comparing AI tutoring vs. no-AI
- Developed *Curio 2*, a **LangGraph**-based science chatbot with scaffolded inquiry, graph-driven dialogue, state management, and **Whisper** for voice interaction.

Volunteer Research Scientist **Sep. 2025 - Present**

Oumi: Open Universal Machine Intelligence *Remote*

- Improving **LLM judges** for better reasoning and reduced inconsistencies; *TrustJudge*, a framework for more reliable evaluations; Analyzing LLM benchmark failures and proposed solutions.

AI/ML Engineer **Sep. 2022 - Jul. 2024**

Tadvin Farayand Co. *Tehran, Iran*

- Analyzed structured ERP and unstructured industrial data; designed AI-driven workflows to enable autonomous collaboration across ERP modules for inventory analysis, document processing, and predictive maintenance.

Selected Publications

[1] Keramati A. and Warschauer M. **"Application of Multi-Agent Systems for Essay Scoring"**, NeurIPS 2025 Education Program (Accepted)

[2] Keramati A., Jie C., Warschauer M. and Shi Y. **"Simulating Students' Java Programming Errors with Large Language Models"**, The International Conference on Learning Analytics & Knowledge (LAK 2026) (Under Review)

[3] Tajik E., Shahrokhian B., Borchers C., Keramati A., Simon S., Pal S. and Sankaranarayanan S. **"Disagreement as Data: Reasoning Trace Analytics in Multi-Agent systems"**, The International Conference on Learning Analytics & Knowledge (LAK 2026) (Under Review)

[4] Keramati A. and Warschauer M. **"The Order Matters: Sequential Fine-Tuning of LLaMA for Coherent Automated Essay Scoring"**, ICLR 2026 (Under Review)

[5] Keramati A. and Warschauer M. **"MA6S: Multi-Agent Supervisor System for Essay Scoring and Feedback Generation with Teacher in the Loop "**, CSCL 2026 (Under Review)

[6] Keramati A., Zhou S., Mehrotra S., and Warschauer M. **"MADRAG as Judge: A Multi-Agent Debate and RAG System for Rubric-Based Evaluation"**, (Pre-print)

[7] Fallah A., Keramati A., Nazari M., and Mirfazeli F. **"Automating Theory of Mind Assessment with a LLaMA-3-Powered Chatbot: Enhancing Faux Pas Detection in Autism"**, *14th International Conference on Computer and Knowledge Engineering (ICCKE 2024)*

[8] Shahhoseini N., Taghiyareh F., and Keramati A. **"iTAG: Easy, Rapid, Automatic Intelligent Tagging for Educational Contents"**, *2025 29th International Computer Conference, Computer Society of Iran (CSICC)*

Technical Skills

Programming Languages: C/C++/C#, Python, Java, JavaScript, R, HTML/CSS, Verilog, Assembly, Stata, SQL
Frameworks & Libraries: LangChain, ReAct, PyTorch, TensorFlow, React, Spring
Technologies & Tools: CUDA, Git, Docker, Kubernetes, Maven, Makefile, Tomcat
Databases: MySQL, Postgres, MongoDB, Neo4j, Redis, ElasticSearch
Operating Systems: Linux, Windows, macOS, Unix
Qualitative Analysis Tools: ATLAS.ti, Dedoose

Honors & Awards

NeurIPS 2025 — AI Education Program **Oct. 2025**
Accepted Oral Presentation *San Diego, CA*

Scholarship Recipient, LearnLab Summer School at Carnegie Mellon University **Jul. 2025**
Awarded funding to attend the program on Computer Science Education Research track at CMU *Pittsburgh, PA*

Graduate Ambassador Candidate – TPC Mentor Fellowship at *UC Irvine* **Jun. 2025**
Chosen to oversee mentor–mentee coordination and selection, enhancing communication and program engagement *Irvine, CA*

Top 10 — B.Sc. Computer Engineering (2019 cohort) at *University of Tehran* **Jul. 2024**
Ranked within the top 10 of the graduating cohort based on cumulative GPA *Tehran, Iran*

Top 0.1%, Konkour — National University Entrance Exam **Aug. 2019**
Ranked among hundreds of thousands of examinees; admitted to B.Sc. in CE at Iran's top-ranked university *Iran*

Selected Research Projects

University of California, Irvine

- *PapyrusAI:* Developed a GPT-based generative writing coach with a **multi-agent RAG pipeline** and autonomous reasoning, enabling to decide when to call tools, functions, or perform web searches. Integrated **fine-tuned BERT** for labeling argumentative essays and **LLaMA-3.1-8B** with **LoRA** + 4-bit **quantization** for essay evaluation.
- Built LLM-powered collaborative classmates with controllable dialogue behaviors; added **GPT-Realtime** voice interaction for **multimodal** problem-solving.
- Analyzed multi-year quiz data to assess the impact of AI usage on student outcomes. **Benchmarked LLMs** on the same assessments to evaluate their proficiency in answering quiz questions.
- Developed an **automated grading and feedback system** for weekly assignments and co-built an **AI detection system** using LLMs to identify AI-generated submissions based on writing patterns, timing, and key concepts.

LearnLab Summer School (2025) at Carnegie Mellon University

- Built a multi-agent **student simulator + verifier** using LLMs to generate Java programs with realistic logical errors. Benchmarked against 74,000+ CodeWorkout submissions and single-agent LLMs. Evaluated error diversity and fidelity with **ASTs** and **ZSS** tree edit distance, using **Input-Output**, **Chain-of-Thought**, and **Self-Refine** prompting strategies.

University of Tehran

- Developed an Open Learner Model integrated with an LLaMA-based personalized tutor within the LMS aiedut.ir; fine-tuned the model via Reinforcement Learning from Human Feedback **(RLHF)** to deliver adaptive feedback for personalized learning.
- Implemented parameter-efficient fine-tuning **(PEFT)** and **RAG** stack with chunking, vector retrieval, re-ranking, and context compression to improve domain-specific reasoning performance.
- Automated generation of interactive learning modules from **IEEE LOM metadata** and integrated Avida-ED simulations; applied genetic algorithms to personalize activity parameters and difficulty levels.
- Built a **LLaMA-3–80B-based chatbot** to administer and evaluate the FPRT assessment through interactive rehabilitation dialogues with adaptive hints, automated scoring (**LLM-as-judge**), and clinician-ready reporting.