Albert Tam

altam@mit.edu ◆ github.com/tam-albert

Education

Massachusetts Institute of Technology *Expected Graduation: May 2026*

B.S. in Computer Science and Engineering, Mathematics GPA: 5.0/5

- **Courses**: Natural Language Processing, Sensorimotor Learning, Computer Systems Engineering, Machine Learning, Computer Architecture, Intermediate Algorithms, Stochastic Processes, Probability and Random Variables, Abstract Algebra

Work Experience

Researcher | **Madry Lab, MIT Computer Science and Artificial Intelligence Laboratory** Jun 2023 – Present

- Researched interpretability methods to attribute image segmentation model predictions to training data.
- Designed dataset curation method that achieves **3x** data efficiency in training robust, generalizable segmentation models.
- Improved quality of popular image datasets by systematically identifying mislabeled training examples.

Software Engineer Intern | **Inkeep** Jan 2024 – Feb 2024

- Enhanced capabilities of a retrieval-augmented generation (RAG) product that uses documentation to answer developer questions.
- Developed feature to comprehend and answer complex user queries, leveraging information retrieval techniques and OpenAI models.
- Contributed to a product with 60,000+ monthly active users, improving answer quality and user satisfaction.

Researcher | **Kellis Lab, MIT Computer Science and Artificial Intelligence Laboratory** Sep 2022 – May 2023

- Developed techniques for factor analysis to integrate multiple single-cell datasets to understand regulatory mechanisms behind complex diseases.
- Used paired PCA methods to locate candidate genetic factors in Alzheimer's disease.

Leadership & Activities

Director | **HackMIT** Sep 2022 – Present

- Leading a team of 35 undergraduates to organize MIT's largest hackathon, with 1,000+ participants annually, in addition to an annual hackathon for 220+ high schoolers.

Projects

Natural Language Terminal Assistant (TreeHacks 2024) Feb 2024

- Built a shell that uses local, fine-tuned Mixtral 8x7B to transform natural language into shell commands in Zig.

Constructing Defenses Against Adversarial LLM Jailbreak Attacks Nov 2023

- Researched prompt tuning for optimizing defenses against published LLM jailbreak attacks to improve model safety.
- Developed pipeline for prompt-tuning and evaluating defenses in PyTorch.

Conversational Language-Learning Chatbot (LAHacks 2023) Apr 2023

- Built tool to help users learn new languages by practicing conversations out loud with a chatbot, using OpenAI and GCP APIs.

Web.lab: Collaborative Journaling App Jan 2023

- Designed and developed journaling website for users to create, share, and collaborate on journals with friends.

Awards

USA Biology Olympiad National Finalist, Top 12 (2021): top 0.5% of high school students. Invited to attend most prestigious biology summer program in the US, and placed top 12 among invited finalists.

American Invitational Mathematics Examination, 5x Qualifier (2018 – 2022): top 5% of high school students.

Publications

Albert Tam, Josh Vendrow, & Aleksander Madry. "Data Attribution for Image Segmentation Models." NeurIPS 2023 Workshop on Attributing Model Behavior at Scale (2023)

Skills

- **Languages/frameworks**: Python, JavaScript, PyTorch, NumPy, React, MongoDB, Express.js, Flask