

Zoey Shu

Fairfax, VA | <https://upsilonv0.vercel.app> | github.com/ZhiyaoShu | zshu2@gmu.edu | (201) 686-3469

Human-AI Collaboration; Computer vision; Multimodal sense-making; Computational social study

Publications

- [pre-print] Jie Gao, Erika Lee, Zhiyao Shu, Michelle Vaccaro, Xiaoxian Zhang, SoeMin, Thomas W Malone. Not Just Prompt Engineering: A Taxonomy for Human-LLM Interaction Modes, 2025
- [pre-print] Jie Gao, Zhiyao Shu, ShunYi Yeo. MindCoder: Supporting Ideation in Inductive Qualitative Analysis with Large Language Models, 2025
- Zhiyao Shu, Xiangguo Sun, Hong Cheng. 2024. When LLM Meets Hypergraph: A Sociological Analysis on Personality via Online Social Networks. In Proceedings of the 33rd ACM International Conference on Information and Knowledge Management (CIKM '24). Association for Computing Machinery, New York, NY, USA, 2087–2096. <https://doi.org/10.1145/3627673.3679646>
- Zhiyao Shu. 2024. Thumb up or Head down? The Impact of Upvote on Social Media Self-efficacy. Lecture Notes in Education Psychology and Public Media, 37, 138–145.

Ongoing Project

- Visual place recognition task: Developed classifiers to train and fine-tuning **DINO-v2 vision transformer** as the backbone on a **40-million-image dataset**.

Experiences

Research Intern

Singapore-MIT Alliance for Research & Technology Centre

Remote | Jun - Dec 2024

- Led the system implementation of a theory-driven AI platform for inductive qualitative analysis. Architected a global, hierarchical data management system with **Zustand**. Optimized data retrieval from user inputs and file uploads, integrating structured prompt embeddings with **OpenAI API**. Leading a **40%** improvement in data fetching speed and reduced **20%** data loss.
- Leveraged **LLamaIndex** and **function calling API** for embedding and structuring data to enhance the utilization of large language models (LLMs). Lead **23%** higher satisfaction of outputs in user interviews.
- Authored key sections of one application-associated research paper, including system architecture, primary features, and related works. Created detailed diagrams for experiment result analysis. The manuscript has been submitted to **User Interface Software and Technology (UIST) 2025** and is currently under review.
- Contributed to human-AI interaction (HAI) collaboration research, annotating over **300** publications from ACL, CSCW, IUI 2021 - 2024 to classify and evaluate existing HAI interaction frameworks. Discuss with the first author and co-authored the manuscript.

Software Engineer

Nexa AI

San Jose, CA | Jan - Jun 2024

- Collaborated with a five-member engineering team to develop a scalable AI-powered React application, utilizing **Context API** for efficient data management. Partnered with two product managers to optimize user flows and build responsive, scalable front-end components.

- Developed the application backend with **Java** and **Spring MVC structure** to manage HTTP requests and responses. Constructed aggregation pipelines using **Spring MongoDB** and **MongoTemplate** to perform querying and manipulation in MongoDB databases.
- Led the website **SEO** enhancement solutions, with effective image scaling and **content delivery network (CDN)** optimization on **GCP**, reducing **30%** reduction (from 126.3ms to 37.62 ms) in loading latency.
- Utilized **llama.cpp** to quantize language models on **Ollama (LangChain)** to convert models to Q2 - F16 **GGUF** formats. Published and developed quantized models on **HuggingFace** Spaces, enabling direct interaction with the updated models and facilitating real-time feedback. Resulting in a **33%** improvement in content generation quality and a 1.5% decrease in memory usage. The quantized models acquired **1521** overall downloads.
- Applied **LoRa** for fine-tuning a language model to facilitate cross-application collaboration tasks.
- Engineered the **mobile version** of the platform using **Java** in **Android Studio**, developing essential utility class modules to enable a function-driven language model supporting on-device actions. Product were demoed for VP

Skills

- **Machine learning:** Linear regression, Logistic regression, Random forest
- **Neural Network:** Large language model (LLM), Sentence-transformer, Graph neural network(GNN), Convolutional neural network (CNN)
- **Data Visualization:** D3.js, Matplotlib, Seaborn, Adobe illustrator
- **Programming:** Python, Javascript/Typescript, React.js, Node.js, MongoDB, HTML/CSS, Java, Next.js
- **App development & Platform:** Github, Docker, GCP, Android Studio, linux

Education and Certification

George Mason University

Ph.D. of Information Science and technology

Fairfax, VA | Aug 2025

MIT Professional Education

Applied Data Science Program: Leveraging AI for Effective Decision-Making

Remote | Jul - Nov 2023

University of California, Berkeley

Master of Landscape Architecture and Environmental Design

Berkeley, CA | Aug 2018 - Jun 2020

Colorado State University

Bachelor of Science

Fort Collins, CO | Aug 2014 - Aug 2018