Zixuan (Arya) Wu

zw102@wellesley.edu • (781) 366-8967 • aryawu0513.github.io • linkedin.com/in/arya-zixuan-wu

EDUCATION

Wellesley College, Wellesley MA
B.A. in Computer Science, Architecture | GPA: 3.96/4.0 **Dec 2024**
Relevant Courses: Data Structure, Algorithms, Databases with Web Interfaces, Mobile App Development, Distributed Computing, Computer System Engineering, Regression Analysis and Statistical Models.

Massachusetts Institute of Technology, Cambridge MA **Dec 2024**
Cross Registered Student **|** GPA: 5.0/5.0
Relevant Courses: Natural Language Processing, Machine Learning, Computer Vision, Computer Graphics, Advanced Interaction Studio, Operating System Engineering.

RESEARCH INTERESTS

Machine Learning, Natural Language Processing, Large Language Models, Multimodal AI, LLM for Code, Software Engineering, Mechanistic Interpretability.

PUBLICATIONS

1. Lucchetti, Francesca, **Zixuan Wu**, Arjun Guha, Molly Q Feldman, and Carolyn Jane Anderson. 2024. Substance Beats Style: Why Beginning Students Fail to Code with LLMs. https://arxiv.org/abs/2410.19792. In Submission to NAACL 2025.
2. **Wu, Zixuan**, Yoolim Kim, and Carolyn Jane Anderson. 2024. GlyphPattern: An Abstract Pattern Recognition for Vision-Language Models. https://arxiv.org/abs/2408.05894. Under review at ACL Rolling Review Dec 2024.
3. **Wu, Zixuan**, Christine Bassem. 2024. Bundle and Route: Spatio-Temporal Clustering and Scheduling of Origin-Destination Pairs for On-Demand Food Delivery. Manuscript.

RESEARCH EXPERIENCE

Northeastern Programming Languages Lab, Boston MA
Pre-Doctoral Research Assistant **Spring 2025**
Student Researcher **May 2024 - Present**
- National Deep Inference Fabric

- Reproduce existing mechanistic interpretability methods using the NDIF NNsight package.
- Engineer code-free interfaces for using logit lens and tuned lens on user data, deployed as Gradio space.
- Code batch generation pipeline that support resumption, supports various base and instruct models and vllm.

- Code LLM Interpretability

- Construct one-token prediction pointer problems benchmark by A-normalizing NetworkX random graph codes with enforced pointer structures of varying complexities.
- Build LLM code output parsing & execution pipeline allowing concurrency and safe subprocess management.
- Evaluate models and apply interpretability tools to model internal computation on these pointer problems.

- Student-LLM Code Prompt Evaluation

- Conduct counterfactual causal intervention experiments by developing a pipeline to tag and substitute non-canonical wording by programming concepts terminology, accounting for morphological variants.
- Engineer an automated system to visualize prompts as graphs of clue editing trajectories.

Wellesley Empirical and Statistical Explorations of Language Lab, Wellesley MA **Jan 2024 - Present**
Student Researcher **| Benchmarking Vision-Language Models on Abstract Reasoning**

- Code data processing, image construction, prompt engineering pipeline to create 954 instances of multiple choice puzzles with three visual presentation styles to evaluate model's spatial and compositional reasoning.
- Evaluate dataset on Vision-Language Models, with zero-shot, few-shot, and chain-of-thought prompting.
- Conduct statistical analysis and qualitative error case evaluations.

MIT Media Lab Viral Communications Group, Cambridge MA **Jan 2024 - Present**
Software Engineer Intern**| LatentLab.ai: Graphical Knowledge Exploration Interface for Visual-enhanced RAG**

- Build upon UMAP dimension reduction of data topics to develop a knowledge graph generator, visualized as topological maps using deck.gl layers, within a frontend developed with React and TypeScript.

- Implement API routes for dataset scraping, loading, and KG construction on App Server in Python FastAPI.
- Design and conduct HCI experiments measuring the effect of augmenting retrieved content with visualizations to promote user-driven exploration.
- Code frontend interfaces to support controlled experimental conditions and implement user interaction metric calculations. Embed these into a designed Qualtrics survey with Google Sheets integration for data logging.

Wellesley CrowdNets Lab **Sep 2023 - June 2024**
Student Researcher | **Crowdsourced food delivery**
- Designed an ML pipeline of customized clustering algorithms on spatial-temporal food delivery data.
- Coded a Java Ride Simulator for agent-based event simulation of offline approximated order batch & match.
- Developed a Greedy Insertion Vehicle Routing Program algorithm for customer fairness, and a Bipartite Matching algorithm for courier fairness. Performed statistical visualization to evaluate the methods.

MIT Ideation Lab **Jan 2023 - May 2023**
Student Researcher | **User Centered Design using AI (DOI: 10.1115/DETC2023-116601)**
- Ran 30 social experiments in Grasshopper on Generative Design Tools in multi-objective optimization design.
- Designed research experiments on the carrying of aesthetic preferences using conjoint analysis.
- Performed statistical analysis tests in R, produced data visualizations and conclusions for paper writing.

SELECTED COURSE PROJECTS

Next AI: Full-Stack Web-Based Q&A Al Agent for PDF Document Queries by React and Node.js
- Created an interactive conversational UI for users to upload and interact with PDF documents in real-time.
- Implemented RESTful APIs via Express and Node.js and optimized for high-performance request handling.
- Utilized an in-memory vector store to cache generated embeddings for efficient retrieval.
- Integrated OpenAI API and Langchain for document loading, splitting, storage, retrieval, and output.

OnlineOrder: A food ordering web app based on Spring Boot, ReactJS, AntDesign
- Developed CRUD REST APIs with Spring Controllers, support registration, menu searching, ordering, checkout.
- Utilized Spring Data JDBC for PostgreSQL database hosted on AWS RDS, handling menus, restaurants data.
- Implemented Spring Security for session-based authentication.
- Built frontend using React and Ant Design, allowing users to add shopping cart and place orders.

Around: A Cloud and React based Social Network
- Implemented a social network web application using Go, where users can create, browse, and search posts.
- Improved the authentication using JWT token based registration/login/logout flow with React Router v4.
- Used ElasticSearch as a NoSQL database, with inverted index for messages to perform keyword-based search.
- Used GCS to store all media files, with links stored as metadata in ElasticSearch.

Humoment: A Large-Language Object that captures day with music
- Built a Software Python code pipeline for music creation, leveraging Microsoft's MusicGen and AudioGen, fine-tuned on the user's playlist for humming tune continuation.
- Built a conversational chatbot using GPT4 whisper, with emotion analysis to synthesize information.
- Designed the product using 3D Modeling, powder printing, house Arduino and Xiao electronic components.

VMetaphor: Automatic Generation of Advertisement Visual Metaphors
- Collected visual metaphor datasets and finetuned Stable Diffusion Model as an advertisement design tool.
- Coded a generative pipeline including parts segmentation, masking, image inpainting edits using pyTorch.
- Conducted Ablation study across models and components, with qualitative & quantitative analysis for writing.

FELLOWSHIP & LEADERSHIP EXPERIENCE

- NCAA Division III Academic All-Conference | Wellesley Varsity Swimming Team 2021 - 2024
- Teaching Assistant | Wellesley Mathematics, Computer Science Department 2022 - 2024
- Volunteering Program Manager | Wellesley Robogals Global 2022 - 2023
- Slade Summer Fellowship awardee| Tanner Conference Speaker | Wellesley Art Department 2022 - 2023

TECHNICAL SKILLS

Programming Languages: Python, Java, JS/TS, C/C++, Golang, R, MATLAB. (in level of proficiency)
Development Tools: HTML/CSS, React, React Native, Node.js, Flask, MySQL, MongoDB, Springboot, PyTorch, OpenCV, Git, Docker, AWS EC2, Figma.