# Zexin (Jason) Xu

zexin.xu@utdallas.edu | zexinxu.com | in/zexin-xu | github.com/asonjay

## Research Interests

Responsible AI, Alignment, Embodied Agent, Deep Learning, Educational Technology

## Education

**The University of Texas at Dallas** — Aug 2024 – Present
*Ph.D. Student in Electrical & Computer Science; Advisor: Dr. Wei Yang* — *GPA: 4.00/4.00*

**The Ohio State University** — Aug 2021 – May 2023
*M.S. in Computer Science and Engineering; Advisor: Dr. Yu Su* — *GPA: 3.51/4.00*

**The Ohio State University** — Jan 2018 – Dec 2020
*B.S. in Computer Science and Engineering / Magna Cum Laude, Dean's List* — *GPA: 3.87/4.00*

## Experience

**Team Lead** — Aug 2024 – Jun 2025
*The University of Texas at Dallas (Amazon Nova AI Challenge: Trusted AI)* — *Richardson, TX*

- Led team ASTRO representing UTD, participated in the first Nova AI Trusted AI challenge on Code LLM red teaming and defense, selected as one of **5 red teams globally**.
- Achieved **2nd place in 2 tournaments** and finalist status; published paper in Amazon Science proceedings.
- Developed red teaming pipeline to trigger defense models on generating malicious/vulnerable code using **CodeGuru, Bedrock, Cobot** and other AWS services.

**Software Engineering Lead** — Nov 2023 – Present
*University of California, Irvine (Part Time)* — *Irvine, CA*

- Co-led TimeWise educational conversational agent project with researchers from **UCI, NTU, and JHU**.
- Selected as **finalist in 2024 Tools Competition** among 2,000 submissions; built React-based Chrome extension with full AWS infrastructure.

**Research Assistant, Software Engineering Lead** — Aug 2022 – Jun 2023
*The Ohio State University (Amazon Alexa Prize SimBot Challenge)* — *Columbus, OH*

- Led engineering team for SalsaBot in the first Alexa Prize Simbot Challenge on **embodied agent**, architecting state-driven dialogue system with language macro module using AWS.
- Elevated team rating from **2.36 to 3.64 out of 5.00**, achieving top-three standing in final month through comprehensive testing methodologies.

**Data Scientist Intern** — May 2022 – Aug 2022
*AI Camp* — *Palo Alto, CA*

- Directed a team of over 15 students in developing cutting-edge NLP web applications, harnessing the power of **GPT-2 and Flask**. This initiative culminated in an outstanding **average rating of 4.53/5.00**.
- Modernized the internal rating mechanism and spearheaded pivotal discussions on curriculum enhancement.
- Played a key role in shaping the NLP, deep learning, and Flask curricula, and actively participated in periodic curriculum updates.

**Software Developer Lead** — Sep 2020 – Dec 2020
*Sponsored by Honda* — *Columbus, OH*

- Designed and implemented an advanced self-service kiosk dialogue system for internal applications, equipped with chatbot functionalities, enhanced by integrating a cutting-edge language understanding module using **Microsoft Azure Bot Framework** and **LUIS**.
- Established a resilient software architecture and streamlined the development workflow.

## Publications

Note: * denotes equal contribution. Full list: Google Scholar

1. **COMET: Closed-loop orchestration for malicious elicitation techniques in code models**
   Zexin Xu, Tingxi Li, Ravishka Shemal Rathnasuriya, Zihe Song, Jun Ren, Bhavesh Mandalapu, Soroush Setayeshpour, Xinya Du, Wei Yang
   *Amazon Nova AI Challenge Proceedings, 2025*

2. **LLM4SR: A survey on large language models for scientific research**
   Ziming Luo*, Zonglin Yang*, Zexin Xu, Wei Yang, Xinya Du
   *arXiv preprint, 2025*

3. **SoK: Efficiency Robustness of Dynamic Deep Learning Systems**
   Ravishka Rathnasuriya, Zexin Xu*, Tingxi Li*, Zihe Song*, Mirazul Haque, Simin Chen, Wei Yang
   *arXiv preprint, 2025*

4. **Beyond pass or fail: A multi-dimensional benchmark for mobile UI navigation**
   Dezhi Ran, Mengzhou Wu, Hao Yu, Yuetong Li, Jun Ren, Yuan Cao, Xia Zeng, Haochuan Lu, Zexin Xu, Mengqian Xu, Ting Su, Liangchao Yao, Ting Xiong, Wei Yang, Yuetang Deng, Assaf Marron, David Harel, Tao Xie
   *arXiv preprint, 2025*

5. **SalsaBot: Towards a Robust and Generalizable Embodied Agent**
   Chan Hee Song, Jiaman Wu, Ju-Seung Byun, Zexin Xu, Vardaan Pahuja, Goonmeet Bajaj, Samuel Stevens, Ziru Chen, Yu Su
   *Embodied AI Workshop at CVPR, 2023*

6. **Towards a Robust and Generalizable Embodied Agent**
   Chan Hee Song, Jiaman Wu, Ju-Seung Byun, Zexin Xu, Vardaan Pahuja, Goonmeet Bajaj, Samuel Stevens, Ziru Chen, Yu Su
   *Amazon Science, 2023*

7. **Exploring the Role of Artificial Intelligence in Facilitating Assessment of Writing Performance in Second Language Learning**
   Zilu Jiang, Zexin Xu, Zilong Pan, Jingwen He, Kui Xie
   *Languages, 2023*

8. **Measuring Elementary Students' Behavioral Engagement in Web-based Science Inquiry Learning**
   Jingwen He, Bihui Jin, Zexin Xu, Danhui Zhang
   *Journal of Online Learning Research, 2022*

## Teaching

**CS4375 Introduction to Machine Learning** — The University of Texas at Dallas — Fall 2025
*Graduate Teaching Assistant*

**CE4337 Programming Language Paradigms** — The University of Texas at Dallas — Fall 2025
*Graduate Teaching Assistant*

**CS5343 Algorithm Analysis and Data Structures** — The University of Texas at Dallas — Summer 2025
*Graduate Teaching Assistant*

**SE3377 Systems Programming** — The University of Texas at Dallas — Summer 2025
*Graduate Teaching Assistant*

**CE4337 Programming Language Paradigms** — The University of Texas at Dallas — Spring 2025
*Graduate Teaching Assistant*

**CE4337 Programming Language Paradigms** — The University of Texas at Dallas Fall 2024
*Graduate Teaching Assistant*

**CSE6521 Advanced Artificial Intelligence** — The Ohio State University Fall 2022
*Graduate Teaching Associate*

**CSE2321 Foundations I: Discrete Structures** — The Ohio State University Spring 2020
*Undergraduate Teaching Assistant*

**CSE3521 Artificial Intelligence I: Basic Techniques** — The Ohio State University Fall 2019
*Undergraduate Teaching Assistant*

## Awards

**Amazon NOVA AI Challenge (Trusted AI) Grant** Jan 2025 - Jul 2025
*Amount Awarded: $250,000*

**Amazon NOVA AI Challenge (Trusted AI) AWS Research Credit** Oct 2024 - Jul 2025
*Amount Awarded: $150,000 per month, Total Amount Awarded: $1,350,000*

## Skills

**Certificate**: AWS Cloud Technical Essentials
**Languages**: Python, Java, C/C++, C#, SQL, JavaScript, R, HTML/CSS
**Frameworks & Tools**: AWS, Django, Flask, Node.js, Next.js, Docker, Microsoft Azure Bot Framework, Git, CUDA
**Libraries**: LangChain, PyTorch, OpenCV, Pandas, NumPy, Scikit-learn, Matplotlib, OpenMP
**Databases**: MongoDB, MySQL, PostgreSQL

## Reference

**Wei Yang, Ph.D.**
The University of Texas at Dallas
*Associate Professor, Department of Computer Science*
**Email**: wei.yang@utdallas.edu; **Website**: youngwei.com

**Yu Su, Ph.D.**
The Ohio State University
*Associate Professor, Department of Computer Science and Engineering*
**Email**: su.809@osu.edu; **Website**: ysu1989.github.io

**Di Xu, Ph.D.**
University of California, Irvine
*Professor, School of Education*
**Email**: dix3@uci.edu; **Website**: faculty/dixu

**Kui Xie, Ph.D.**
University of Missouri, Columbia
*Professor, Dean and Joanne H. Hook Dean's Chair for the College of Education & Human Development*
**Email**: xiekui@missouri.edu; **Website**: person/kui-xie

**Zilong Pan, Ph.D.**
Lehigh University
*Assistant Professor of Teaching, Learning and Technology*
**Email**: zip322@lehigh.edu; **Website**: ed.lehigh.edu/faculty/zpan

**Zilu Jiang, Ph.D.**
Johns Hopkins University
*Postdoc Fellow of School of Education*
**Email**: zjiang72@jh.edu