

Zeyu (Steven) He 何泽宇

Ph.D. Student
College of Information Sciences and Technology
Pennsylvania State University
E393 Westgate Building, University Park, PA 16802

Email: zmh5268@psu.edu
zeyuhe98@gmail.com
Homepage: <https://stevenhe918.github.io/>

Research Areas

My research explores Human-Computer Interaction (HCI) and Natural Language Processing (NLP), focusing on human involvement in the data annotation pipeline in the era of Large Language Models (LLMs).

Education

- **Ph.D., Informatics**
College of Information Sciences and Technology, Pennsylvania State University, 2022 – Present
Advisor: Dr. Kenneth Huang.
- **B.S., Computer Science and Engineering**
Henry Samueli School of Engineering, University of California Irvine, 2016 – 2020
GPA: 3.74/4.0
- **B.S., Civil Engineering (Double Major)**
Henry Samueli School of Engineering, University of California Irvine, 2016 – 2020
GPA: 3.74/4.0

Experience

- **Pennsylvania State University**, University Park, PA, USA, 08/2022 – Present.
Ph.D. Research Assistant, Advisor: Dr. Kenneth Huang.
- **Amazon (Amazon Music Recommendation Service Team)**, Irvine, CA, USA, 02/2022 – 08/2022.
Software Development Engineer (Full-Time)
Developed front and back-end support for the recommendation service, enhancing the user experience.
- **Personable**, Fountain Valley, CA, USA, 08/2020 – 02/2022.
Software Engineer (Full-Time)
Delivered financial software solution for image processing, data extraction, and data management.

Awards and Honors

- AAAI HCOMP Doctoral Consortium Financial Support (\$400), October 2024
- Penn State IST Travel Award (\$1500), May 2024
- UC Irvine Dean's Honor List, 2016-2020

Peer-Reviewed Conference Papers

- [C.4] **Zeyu He**, Saniya Naphade, Ting-Hao K. Huang. (2025). Prompting in the Dark: Assessing Human Performance in Prompt Engineering for Data Labeling When Gold Labels Are Absent. In Proceedings of Conference on Human Factors in Computing Systems 2025 (**CHI 2025**) (Under Review)
- [C.3] Ho Yin Ng, **Zeyu He**, Ting-Hao K. Huang. (2024). What Color Scheme is More Effective in Assisting Readers to Locate Information in a Color-Coded Article? IEEE Transactions on Visualization and Computer Graphics (**VIS 2024**), October, 2024, Tempa, FL, USA.
- [C.2] **Zeyu He**, Chieh-Yang Hunag, Chien-Kuang (Cornelia) Ding, Shaurya Rohatg, Ting-Hao K. Huang. (2024). If in a Crowdsourced Data Annotation Pipeline, a GPT-4. In Proceedings of Conference on Human Factors in Computing Systems 2024 (**CHI 2024**), May, 2024, Honolulu, Hawaii, USA. *Invited to present at CHI2024 Workshop on LLMs as Research Tools (L*ART): Applications and Evaluations in HCI Data Work*
- [C.1] Jakob Hederich, Shreya Ghosh, **Zeyu He**, Prasenjit Mitra. (2023). Understanding the Night-Sky? Developing AI-Enabled System for Exploring Night-Light Usage Patterns. The 32nd International Joint Conference on Artificial Intelligence (**IJCAI 2023**), October, 2023, Macao, China.

Peer-Reviewed Journal Articles

- [J.1] Lyuyang Hu, Omkar Pathak, **Zeyu He**, Hunkyu Lee, Mina Bedwany, Jace Mica, Peter J Burke. (2020). "CloudStation:" A Cloud-Based Ground Control Station for Drones. IEEE Journal on Miniaturization for Air and Space Systems (2020). (Published)

Workshop, Symposia, and Consortia Papers

- [W.1] **Zeyu He**. Human Involvements in Data Annotation in the Era of Large Language Models. (**HCOMP DC 2024**), October, 2024, Pittsburgh, PA, USA.

Posters and Extended Abstracts

- [A.1] Shih-Hong Huang, Ya-Fang Lin, **Zeyu He**, Chieh-Yang Huang, Ting-Hao K. Huang. (2024). How Does Conversation Length Impact User's Satisfaction? A Case Study of Length-Controlled Conversations with LLM-Powered Chatbots. In CHI '24 Late-Breaking Work on Human Factors in Computing Systems (**CHI LBW 2024**), May, 2024, Honolulu, Hawaii, USA.

Professional Activity

- **Paper Reviewer:** CHI LBW 2024, Learning @ Scale 2024, NeurIPS Workshop 2024
- **Treasurer:** Graduate Student Association in Information Sciences and Technology

Skills

- **Programming and Scripting Languages:** Python, Java, C++, R, SQL, Kotlin, C, Verilog, Visual Basic, Matlab, HTML, CSS, JavaScript, TypeScript, PHP
- **NLP, Machine Learning, and Statistics:** GPT, Hugging Face, TensorFlow, PyTorch, Spacy, Gensim, NLTK, BERT-based models, transformer models, classification, clustering, regression, hypothesis testing, Bayesian analysis, data visualization
- **Crowdsourcing and Human Computation:** Designing and managing crowdsourced tasks, quality control in crowd work, human-in-the-loop systems, data collection, and aggregation from crowd-workers.
- **Interface Design, Prototyping, and Front-end Development:** Premier Pro, Figma, Photoshop, Sketch, storyboarding, React

References

- **Kenneth Huang**
Associate Professor, College of Information Sciences and Technology, Pennsylvania State University, USA.
txh710@psu.edu

Last updated: October 25, 2024