

YUXUAN LU

✉ lu.yuxuan@northeastern.edu ·  yuxuan-lu ·  https://yuxuan.lu · 📞 781-870-7184 ·  

🔍 RESEARCH FOCUS

- The intersection between NLP & HCI, “Human-Centered NLP”
- LLM-Agent, Human-AI collaboration, Data Annotation
- Utilizing LLMs efficiently and effectively

🎓 EDUCATION

Northeastern University, MA, USA

Sep. 2023 – Present

Ph.D. in Computer Science. Advised by Prof. Dakuo Wang

Beijing University of Technology, Beijing, China

Sep. 2019 – Jun. 2023

Graduated **with honor**. Bachelor of Engineering in Computer Science and Technology

🏆 HONORS AND AWARDS

🏆 NSF Access Grant 2024

🏆 3 × *Bronze Medal*, ICPC Asia Regional Contest 2019, 2020, 2021

🏆 *Global Rank 42(Top 2%) and 85 (Top 3.5%)*, IEEEExtreme programming competition 2020, 2021

📖 PUBLICATIONS

- [1] **Yuxuan Lu**, Bingsheng Yao, Shao Zhang, Yisi Sang, Yun Wang, Hansu Gu, Peng Zhang, Tun Lu, Toby Jia-Jun Li, and Dakuo Wang. “ALERTS: Active Learning and Ensemble LLM Real-Time Switch for Real-World Data Drift Challenges”. In: *Submission to EMNLP 2024*. 2024.
- [2] Jiaju Chen, **Yuxuan Lu**, Shao Zhang, Bingsheng Yao, Yuanzhe Dong, Ying Xu, Yunyao Li, Qianwen Wang, Dakuo Wang, and Yuling Sun. “StorySpark: Expert-Annotated QA Pairs with Real-World Knowledge for Children Storytelling”. In: *Submission to EMNLP 2024*. 2024.
- [3] Jiaju Chen, **Yuxuan Lu**, Shao Zhang, Bingsheng Yao, Yuanzhe Dong, Ying Xu, Yunyao Li, Qianwen Wang, Dakuo Wang, and Yuling Sun. “Exploring Domain Adaptation with LLMs for Real-World Augmented Question Answer Generation (RA-QAG) in Children Storytelling”. In: *Submission to EMNLP 2024*. 2024.
- [4] Ziqi Yang*, **Yuxuan Lu***, Jennifer Bagdasarian, Vedant Das Swain, Ritu Agarwal, Collin Campbell, Wad-dah Al-Refaire, Dr Jehan El-Bayoumi, Guodong (Gordon) Gao, Nawar Shara, Dakuo Wang, and Bingsheng Yao. “RECOVER: A Large Language Model-based Remote Patient Monitoring System for Postoperative GI Cancer Patients”. In: *Submission to IMWUT 2024*. 2024.
- [5] **Yuxuan Lu**, Bingsheng Yao, Shao Zhang, Yun Wang, Peng Zhang, Tun Lu, Toby Jia-Jun Li, and Dakuo Wang. “Human Still Wins over LLM: An Empirical Study of Active Learning on Domain-Specific Annotation Tasks”. In: *arXiv preprint arXiv:2311.09825* (2023).
- [6] Jiaju Chen, **Yuxuan Lu**, Bingsheng Yao, Qianwen Wang, Ying Xu, Dakuo Wang, and Yuling Sun. “SciSpark: An Interactive Storytelling System for Young Children’s Science Education with Large Language Model (LLM) and Retrieval-Augmented Generation (RAG)”. In: *Submission to UIST 2024*. 2024.
- [7] Shao Zhang, Shihan Fu, Bin Lu, **Yuxuan Lu**, Toby Jia-Jun Li, Dakuo Wang, Ying Wen, Xinbing Wang, and Chenghu Zhou. “From Dark Data to Open Data: Challenges and Practices for Data Integrators of Data-Driven Open Science Projects in Geoscience”. In: *Submission to CSCW 2025*. 2024.

- [8] Shao Zhang, Jianing Yu, Xuhai Xu, Changchang Yin, **Yuxuan Lu**, Bingsheng Yao, Melanie Tory, Lace M. Padilla, Jeffrey Caterino, Ping Zhang, and Dakuo Wang. “Rethinking Human-AI Collaboration in Complex Medical Decision Making: A Case Study in Sepsis Diagnosis”. In: *Proceedings of the CHI Conference on Human Factors in Computing Systems*. CHI ’24. New York, NY, USA: Association for Computing Machinery, 2024. DOI: 10.1145/3613904.3642343.
- [9] Yuling Sun, Jiali Liu, Bingsheng Yao, Jiaju Chen, Dakuo Wang, Xiaojuan Ma, **Yuxuan Lu**, Ying Xu, and Liang He. “Exploring Parent’s Needs for Children-Centered AI to Support Preschoolers’ Storytelling and Reading Activities”. In: *Proc. ACM Hum.-Comput. Interact.* 7.CSCW2 (2024). URL: <https://arxiv.org/abs/2401.13804>.
- [10] Bingsheng Yao, Guiming Chen, Ruishi Zou, **Yuxuan Lu**, Jiachen Li, Shao Zhang, Sijia Liu, James Hendler, and Dakuo Wang. “More Samples or More Prompt Inputs? Exploring Effective In-Context Sampling for LLM Few-Shot Prompt Engineering”. In: *Findings of the Association for Computational Linguistics: NAACL 2024*. 2024.
- [11] Bingsheng Yao, Ishan Jindal, Lucian Popa, Yannis Katsis, Sayan Ghosh, Lihong He, **Yuxuan Lu**, Shashank Srivastava, Yunyao Li, James Hendler, and Dakuo Wang. “Beyond Labels: Empowering Human Annotators with Natural Language Explanations through a Novel Active-Learning Architecture”. In: *Findings of the Association for Computational Linguistics: EMNLP 2023*. Association for Computational Linguistics, Dec. 2023.
- [12] Hao Chen, Lun Du, **Yuxuan Lu**, Qiang Fu, Xu Chen, Shi Han, Yanbin Kang, Guangming Lu, and Zi Li. “Professional Network Matters: Connections Empower Person-Job Fit”. In: *Proceedings of the 17th ACM International Conference on Web Search and Data Mining*. WSDM ’24. New York, NY, USA: Association for Computing Machinery, 2024.
- [13] **Yuxuan Lu**, Jingya Yan, Zhixuan Qi, Zhongzheng Ge, and Yongping Du. “Contextual Embedding and Model Weighting by Fusing Domain Knowledge on Biomedical Question Answering”. In: *Proceedings of the 13th ACM International Conference on Bioinformatics, Computational Biology and Health Informatics*. BCB ’22. New York, NY, USA: Association for Computing Machinery, 2022. DOI: 10.1145/3535508.3545508.

WORKING EXPERIENCES

Amazon Palo Alto, CA, USA

Sep. 2024 – Present

Intern Applied Scientist Manager: Jing Huang

- Shopper Action Model: Fine-tune a LLM for accurate simulation of user behavior in e-commerce web platforms.

LinkedIn & Microsoft Research Asia Beijing, China

Jul. 2022 – May. 2023

Machine Learning Researcher Manager: Guangming Lu (Manager) / Lun Du (Sr. Researcher)

Participated in a joint program between Microsoft Research Asia and LinkedIn, to discover AI applications driven by LinkedIn’s large-scale high-quality production data.

- Training framework for PyTorch models in LinkedIn’s internal Kubernetes cluster, contributed several patches on HDFS support for upstream packages (fsspec, pytorch_lightning, ...)
- Heterogeneous Knowledge-based Person-Job Fit, **accepted to WSDM 2024 [12]**

SKILLS

- Programming: Multilingual. Fluent in Python, C++, Rust, Go, JavaScript, etc.
- Deep Learning Framework: PyTorch, Hands-on experience in large-scale parallel training
- Data processing: Apache Spark, Pandas
- Languages: English - Fluent (TOEFL 108), Mandarin - Native speaker

SERVICES

Lead Organizer, CS PhD Admissions Committee @ NEU	2023
Reviewer @ EMNLP	2023, 2024
Reviewer @ NAACL	2024
Reviewer @ CHI	2024