YISI SANG

yisisang@gmail.com ◇ Google Scholar, Boston, MA

SUMMARY

I am a researcher and machine learning engineer with an interdisciplinary background in Natural Language Processing (NLP), Human-Computer Interaction (HCI), and Psychometrics (focusing on evaluation and measurement). My previous experience in human-in-the-loop task design, large-scale data pipeline products, large language model (LLM) research and evaluation, and user interface/user experience (UI/UX) design equips me with diverse skills, collaborative mindset, and deliverable-driven attitude toward the next exciting adventure in the era of AI. My work has been published in HCI venues such as CSCW, CHI, and NLP conferences such as ACL, NAACL, and EMNLP.

EDUCATION

Ph.D. in Information Science and Technology, School of Information Science, Syracuse University 2022
Dissertation: *Algorithmic Benchmarking and Modeling of Character Understanding in Narrative Stories*

M.S. in Teaching & Curriculum (Quantitative Research, Evaluation and Measurement), School of Education, Syracuse University 2017

B.S. in German, Minor in Psychology School of Foreign Languages, Northwestern Polytechnical University 2015

SKILLS

Programming & Framework	Python, Linux, JavaScript, HTML, PyTorch
Data Analysis & Tool	SQL, Spark, Scala, R, SPSS, NVivo, Figma, AWS

EXPERIENCE

Machine Learning Engineer Aug 2022 - now
Apple, Inc. *Seattle, WA*

- Built knowledge graph utilizing Scala, Spark, and Python.
- Designed and implemented a large-scale data pipeline to evaluate and monitor the quality of entity linking.
- Designed and developed an interactive tool leveraging a knowledge graph and LLMs (Large Language Model) to assess the accuracy and groundness of content generated by LLMs.

Data Scientist Intern Jun 2022 - Aug 2022
Fidelity Investments *Boston, MA*

- Engineered a dialog summarization model, leading to a substantial enhancement in customer service efficiency by extracting concise and pertinent information.

Part-time Lecture May 2022 - Jun 2022
Northeastern University *Boston, MA*

- Instructed the course *ITC6410 Fundamentals of Human Behaviors for Interactive Systems*, providing comprehensive guidance on the interplay between human behaviors and the development of interactive technologies.

GEM Research Volunteer Oct 2021 - Aug 2022
Hugging Face

- Designed and implimented human evaluation criteria and an annotation interface for assessing machine-generated summarizations.

PAPERS & PUBLICATIONS

[1] Yao, B., Chen, G., Zou, R., Lu, Y., Li, J., Zhang, S., **Sang, Y.**, Liu, S, Hendler, J., Wang, D. (2024). More Samples or More Prompts? Exploring Effective Few-Shot In-Context Learning for LLMs with In-Context Sampling. *Proceedings of EMNLP*

[2] Yu, M., **Sang, Y.**, Pu, K., Wei, Z., Wang, H., Li, J., ... & Zhou, J. (2023). Few-Shot Character Understanding in Movies as an Assessment to Meta-Learning of Theory-of-Mind. In submission

[3] Farima Fatahi Bayat., et al. (2023) FLEEK: Factual Error Detection and Correction with Evidence Retrieved from External Knowledge. In *Proceedings of EMNLP-Demos 2023*

[4] Wu, Q., **Sang, Y.**, Wang, D., & Lu, Z. (2023). Malicious Selling Strategies in Livestream E-commerce: A Case Study of Alibaba's Taobao and ByteDance's TikTok. *ACM Transactions on Computer-Human Interaction, 30(3), 1-29. (TOCHI)*

[5] **Sang, Y.** , Mou, X., Yu, M., Li, J., & Stanton, J., (2022) TVSHOWGUESS: Character Comprehension in Stories as Speaker Guessing. In *Proceedings of NAACL 2022*

[6] **Sang, Y.**, Mou, X., Yu, M., Wang, D., Li, J., & Stanton, J. (2022). MBTI personality prediction for fictional characters using movie scripts. In *Proceedings of EMNLP 2022*

[7] **Sang, Y.**, Mou, X., Li, J., Stanton, J., & Yu, M. (2022). A survey of machine narrative reading comprehension assessments. In *Proceedings of IJCAI-ECAI2022*

[8] Xu, Y., et al. (2022). Fantastic Questions and Where to Find Them: FairytaleQA— An Authentic Dataset for Narrative Comprehension. In *Proceedings of ACL 2022*

[9] **Sang, Y.** , & Stanton, J. (2022). The Origin and Value of Disagreement Among Data Labelers: A Case Study of Individual Differences in Hate Speech Annotation. In *International Conference on Information (pp. 425-444). Springer, Cham.*

[10] Gehrmann, S., Bhattacharjee, A., Mahendiran, A., Wang, A., Papangelis, A., Madaan, A., ... & Hou, Y. (2022). Gemv2: Multilingual nlg benchmarking in a single line of code. In *Proceedings of EMNLP-Demos 2022*

[11] **Sang, Y.**, & Stanton, J. (2020). Analyzing Hate Speech with Incel-Hunters' Critiques. In *International Conference on Social Media and Society (pp. 5-13).*

[12] Huang, Y., **Sang, Y.**, Wu, Q., & Yao, Y. (2019). Higher Education Check-Ins: Exploring the User Experience of Hybrid Location Sensing. In *Proceedings of the ACM on Human-Computer Interaction (CSCW).*

[13] Huang, Y., **Sang, Y.**, Wu, Q., & Yao, Y. (2019). Studying User Experience of a Hybrid Location Sensing System. In *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems.*

[14] **Sang, Y.**, Huang, Y. (2019). Understanding Digital Wellbeing Systems through the Lens of the Theory of Planned Behavior (TPB). In *Proceedings of ACM CHI conference*

[15] Wu, Q., **Sang, Y.**, Zhang, Shan, & Huang, Y. (2018). Danmaku vs. Forum Comments: Understanding User Participation and Knowledge Sharing in Online Videos. In *Proceedings of the 2018 ACM Conference on Supporting Groupwork*

TEACHING

- ITC6410 Fundamentals of Human Behaviors for Interactive Systems, Northeastern University
- IST736 Text Mining, Syracuse University
- IST687 Introduction to Data Science, Syracuse University
- IST649 Human-Computer Interactions (HCI), Syracuse University
- EDU647 Statistical Thinking & Application, Syracuse University

PSYCHOMETRICS

- WISC-IV Chinese V ersion Administrator (Intelligence Assessment)
- ABAS-II Chinese Version Administrator (Behavior Assessment)
- Autism Diagnostic Observation Schedule (ADOS-2)