

Contact Information	756 W Peachtree St NW CODA Gatech, Atlanta, GA Tel: 206-306-5850	Homepage: https://xuhuizhou.github.io/ Linkedin: https://www.linkedin.com/xuhuizhou ✉ E-mail: zhouxuhui2018@gmail.com
Research Interests	<ul style="list-style-type: none"> • Natural Language Processing: Socially intelligent and equitable NLP, explainable NLP. • Machine Learning: representation learning, adversarial robustness. 	
Education	<p>Carnegie Mellon University, Pittsburgh, PA Aug 2022</p> <ul style="list-style-type: none"> • PhD in Computer Science (Language Technologies). <p>Georgia Institute of Technology, Atlanta, GA Sep 2021-Aug 2022</p> <ul style="list-style-type: none"> • PhD in Computer Science. <p>University of Washington, Seattle, WA Sep 2019-Jun 2021</p> <ul style="list-style-type: none"> • M.Sc in Computational Linguistics. • Advisor: Noah Smith <p>Nanjing University, Nanjing, China Sep 2015-Jun 2019</p> <ul style="list-style-type: none"> • B.Sc in Statistics, Department of Mathematics • Advisor: Shujian Huang <p>University of California Berkeley (visiting student), Berkeley, CA Aug 2017-May 2018</p>	
Publications (* Equal contribution)	<p>Annotators with Attitudes: How Annotator Beliefs And Identities Bias Toxic Language Detection Maarten Sap, Swabha Swayamdipta, Laura Vianna, Xuhui Zhou, Yejin Choi, Noah A. Smith In <i>Proceedings of NACCL, 2022</i></p> <p>Emergent Communication Fine-tuning (EC-FT) for Pretrained Language Models Shane Steinert-Threlkeld, Xuhui Zhou, Zeyu Liu, C. M. Downey In <i>Proceedings of ICLR EmeCom, 2022</i> 🏆 Runner-up Best Paper</p> <p>Extracting and Inferring Personal Attributes from Dialogue Zhilin Wang, Xuhui Zhou, Rik Koncel-Kedziorski, Alex Marin, Fei Xia In <i>Proceedings of ACL ConvAI, 2022</i></p> <p>Challenges in Automated Debiasing for Toxic Language Detection Xuhui Zhou, Maarten Sap, Swabha Swayamdipta, Noah A. Smith, Yejin Choi In <i>Proceedings of EACL, 2021</i></p> <p>Linguistically-Informed Transformations (LIT): A Method for Automatically Generating Contrast Sets Chuanrong Li*, Lin Shengshuo*, Zeyu Liu*, Xinyi Wu*, Xuhui Zhou* and Shane Steinert-Threlkeld In <i>Proceedings of EMNLP BlackboxNLP, 2020</i></p> <p>Multilevel Text Alignment with Cross-Document Attention Xuhui Zhou, Nikolaos Pappas, Noah A. Smith In <i>Proceedings of EMNLP, 2020</i></p> <p>Evaluating Commonsense in Pre-trained Language Models Xuhui Zhou, Yue Zhang, Leyang Cui, Dandan Huang In <i>Proceedings of AAAI, 2020</i></p> <p>RPD: A Distance Function Between Word Embeddings Xuhui Zhou, Zaixiang Zheng, Shujian Huang In <i>Proceedings of ACL Student Research Workshop, 2020</i></p>	

Manuscripts	Evaluating Gender Bias in Translation between Gendered Languages Xuhui Zhou , Sarthak Garg, Lin Li Submitted to <i>ACL, 2022</i>	
Research Experience	<ul style="list-style-type: none"> • Machine Intelligence @ Apple March 2021–Sep 2021 Research Intern Mentors: Sarthak Garg, Shayne Longpre, Joachim Daiber 	
	<ul style="list-style-type: none"> • Noah’s Ark @ University of Washington Sep 2019–June 2021 Research assistant Mentors: Maarten Sap, Swabha Swayamdipta, Nikolaos Pappas 	
	<ul style="list-style-type: none"> • CLMBR @ University of Washington Jan 2019–Now Research assistant Mentor: Shane Steinert-Threlkeld 	
	<ul style="list-style-type: none"> • Text Intelligence Lab @ Westlake University May 2019–Sep 2019 Research assistant Mentor: Yue Zhang 	
	<ul style="list-style-type: none"> • SUTD NLP @ Singapore University of Technology and Design Jun 2018–Aug 2018 Research intern Mentor: Yue Zhang 	
	<ul style="list-style-type: none"> • DSS @ University of California, Berkeley Jun 2017–Apr 2018 Research committee member 	
Technical Skills	<ul style="list-style-type: none"> • <i>Programming Languages</i>: Python, C/C++, Java, Bash, R • <i>Machine Learning Toolkits</i>: PyTorch, Tensorflow, AllenNLP, Hugging Face Transformers and Datasets, scikit-learn 	
Service	<ul style="list-style-type: none"> • Volunteer: EACL 2021, EMNLP 2020, ACL2020, 2018 Singapore Symposium on NLP • Program Committee & Reviewing: ACL 2021-2022, NLP4PI 2021 	