



Carnegie Mellon University
Language Technologies Institute



Wei Xu
Georgia Institute of Technology

Wei Xu is an assistant professor in the School of Interactive Computing at the Georgia Institute of Technology. Before joining Georgia Tech, she was an assistant professor at Ohio State University since 2016. Xu's research interests are in natural language processing, machine learning, and social media. Her recent work focuses on text generation, semantics, information extraction, and reading assistive technology. She has received the NSF CRII Award, Best Paper Award at COLING, CrowdFlower AI for Everyone Award, and Criteo Faculty Research Award.

The Importance of Data and Linguistics In Neural Language Generation

Natural language generation has become one of the fastest-growing areas in NLP and a popular playground for studying deep learning techniques. Many variants of sequence-to-sequence models with complicated components have been developed. Yet, as I will demonstrate in this talk, creating high-quality training data and injecting linguistic knowledge can lead to significant performance improvements that overshadow gains from many of these model variants. I will present two recent works on text simplification, a task that requires both lexical and syntactic paraphrasing to improve text accessibility: 1) a neural conditional random field (CRF) based semantic model to create parallel training data; 2) a controllable text generation approach that incorporates syntax through pairwise ranking and data argumentation. I will conclude with a brief overview of my group's recent work, including cross-lingual transfer learning, code and entity recognition for StackOverflow.

Friday, November 13, 2020

1:30 - 2:50 PM EST

Join the meeting on Zoom

Meeting ID 968 6753 2227 Passcode 883155

LTI Colloquium 2020-21
is generously sponsored by

abridge