**Text as Data for Psychological Research:**

**Unsupervised Machine Learning Models and Organizational Diversity Research**

Dr. Wei Wang

May 6, 2022

**Abstract**

Text data are ubiquitous in Psychology, yet they are significantly under-appreciated in our research, largely due to the lack of proper methodological guidelines and theoretical frameworks. This presentation introduces an Exploration-Measurement-Validation approach to psychological research with text data and demonstrates applications of this approach in organizational psychology. Specifically, utilizing structural topic modeling (STM)—an unsupervised machine learning method for text mining, this presentation analyzes hundreds of diversity statements and millions of employees’ online ratings on organizational diversity & equity. It is found that the latent semantic topics underlying corporate diversity statements predict employees’ online ratings on organizational effectiveness, shedding light on important theoretical implications for current diversity research and offering practical recommendations for effective diversity management.