

# Visualization of Text Annotations for Corpus Development

Authors: Huan He, Sunyang Fu, Liwei Wang, Andrew Wen, Sijia Liu, Sungrim Moon, Kurt Miller, Hongfang Liu

## Summary:

A gold standard annotated corpus is usually indispensable for building high-quality language models for clinical natural language processing (NLP) tasks. Existing text annotation tools can provide powerful features to cover various needs of text annotation during corpus development, but few tools provide real-time visualization to support the needs of annotation analysis during the annotation process. To address this need, we developed a corpus visualization module in MedTator, a serverless annotation tool.

ACM Author Affiliations: Huan He: Mayo Clinic; Sunyang Fu: Mayo Clinic; Liwei Wang: Mayo Clinic; Andrew Wen: Mayo Clinic; Sijia Liu: Mayo Clinic; Sungrim Moon: Mayo Clinic; Kurt Miller: Mayo Clinic; Hongfang Liu: Mayo Clinic