Differential Privacy for Textual Data

Dina El Zein

Problem: Share useful information about text without sharing private details.

Method: Share embeddings from Large Language Models without sharing the text.

Evaluation: Measure the privacy guarantees using Rényi differential privacy.

Applications: For privacy preserving machine learning applications, i.e. health care.

idiap RESEARCH INSTITUTE