**Leveraging Hierarchical Representations for Preserving Privacy and Utility in Text**

In Machine Learning (ML) tasks and Artificial Intelligence (AI) systems, training data often consists of information collected from users. This data can be sensitive; for example, in conversational systems, a user can explicitly or implicitly disclose their identity or some personal preference during their voice interactions. Explicit personally identifiable information (PII) (such as an individual’s PIN or SSN) can potentially be filtered out via rules or pattern matching.

