**Slide 1: Introduction**

* **Background:** Large language models (LLMs) are increasingly being used in various applications, including biomedical text processing.
* **Problem:** LLMs can memorize and leak their training data, posing privacy risks.
* **Research Question:** Can we apply training data extraction techniques to a smaller, domain-specific LLM like PubMedBERT?
* **Objective:** To investigate the extent to which PubMedBERT memorizes and leaks its training data.

**Slide 2: Methodology**

* **Model:** PubMedBERT, a domain-specific LLM trained on biomedical text.
* **Training Data:** The publicly available PubMedBERT training data.
* **Approach:**
  + Generate text from PubMedBERT using various sampling techniques.
  + Apply membership inference techniques to identify potentially memorized samples.
  + Compare the identified samples to the training data to confirm memorization.

**Slide 3: Expected Outcomes and Significance**

* **Expected Outcomes:**
  + Identify and quantify the amount of training data memorized by PubMedBERT.
  + Analyze the types of data that are most susceptible to memorization.
  + Evaluate the effectiveness of different sampling and membership inference techniques.
* **Significance:**
  + Enhance understanding of memorization in domain-specific LLMs.
  + Raise awareness of privacy risks associated with LLMs in biomedical applications.
  + Inform the development of mitigation strategies to prevent data leakage.