

研究目的

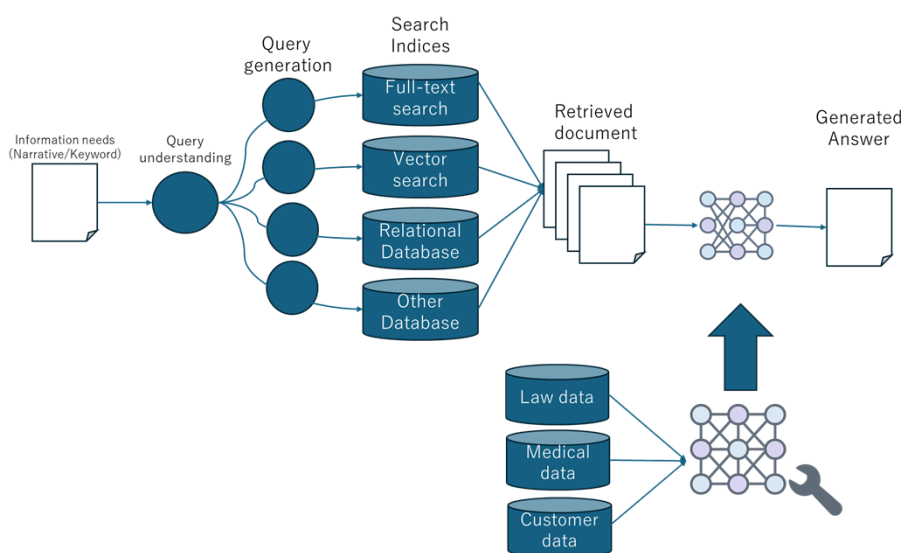
生成AI時代において、情報検索が再度大きな注目を集めています。これは、法律、医療、社内データなど専門的/社内に閉じたデータを検索結果から抽出し、生成AIに入力し回答を得る技術(RAG: Retrieval Augmented Generation)が必要になってきているからです。

本研究室では、実データを用いて、高精度で、効率的に情報検索を行うための研究開発を行っています。

研究概要

本研究室では、大きく2種類の研究を行っています。

1. 生成AIのコンテキストとして入力する情報検索の精度向上に関する研究
2. 特化した分野に対応するための生成AIの精度向上に関する研究



従来・競合との比較

- ・分野特化/専門的なデータを入力とした回答生成を実現
- ・各工程で評価することでどの部分が精度のボトルネックになるのかを確認
- ・医療、Web、法令、製品カタログ等幅広い分野への適用実績

想定される用途

- ・独自のデータを生成AIに活用
- ・現状生成AIを使っても精度が出ない、ボトルネックの特定が困難
- ・数値と文字列等複数の条件でデータを絞り込んだ生成AIの活用

実用化に向けた課題

- ・現状の生成AIのボトルネックである生成に要する速度向上
- ・生成時間と精度のトレードオフについての改善、この点の実用化への課題解決
- ・回答に適したGPU環境、もしくは既存のWebAPIを活用する場合のコスト、精度のトレードオフ解決

企業へ期待すること

- ・分野/に特化したデータと課題を持っている企業
- ・次世代のハードウェアや処理環境を持っている企業
- ・法令等のデータと自社データを組み合わせた生成AIの活用を目指している企業

POINT

- 複数のデータベースの特性を生かした検索
- GPU環境を使った生成AIのファインチューニング

今後の展開

- 2025年 研究室立ち上げ
製造業との共同研究(開始済)
他企業や医療分野との連携(予定)
- 2026年 法令に関するテーマセッション開催(NLP2026)
介護福祉分野への拡大
- 2027年 独自のデータベースや生成AIのモデル公開