

# AIを用いたセマンティック検索システムに関する研究

Research on AI-based Semantic Search Systems

産業システム部 全 慶樹・近藤 正一  
ものづくり支援センター 堀 武司

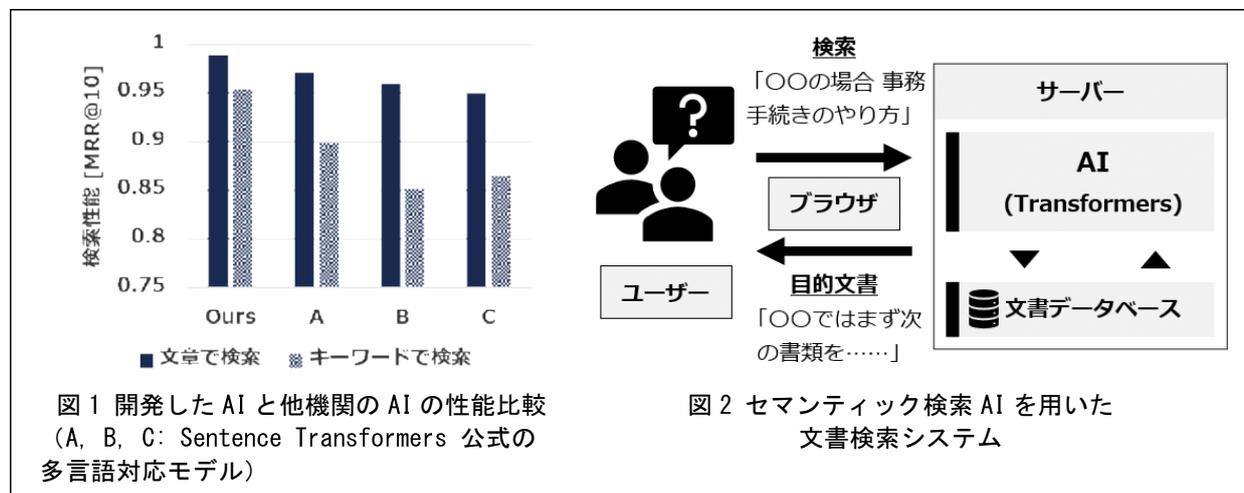
## ■研究の背景

近年、テキストデータを対象とした新しいAI技術が提案され、従来技術では難しかった高度な意味理解に基づく情報検索や質問応答が可能であることがわかってきました。事務作業の多くは、書類やメールといったテキストデータを中心に行われているため、これらのデータにAIを適用することで検索や分類、要約などの作業を効率化できると期待されています。意味に基づく検索はセマンティック検索と呼ばれ、例えば、検索キーワード「スマホ」に対して「携帯電話」や「smartphone」といった類義語や同義語の検索が可能であるほか、キーワードではなく文章による検索が可能であるため、問い合わせ対応を行うヘルプデスク業務の自動化に活用できます。しかし、日本語を対象とした研究事例はまだ少ない状況にあります。

そこで本研究では、企業や自治体の業務効率化支援に向けて日本語を対象としたAIによるセマンティック検索が可能な文書検索システムの開発を行いました。

## ■研究の要点

1. 日本語を対象としたセマンティック検索AIの開発
2. 開発したAIと他機関から公開されているAIの検索性能の評価
3. 開発したAIを用いた文書検索システムの開発



## ■研究の成果

1. 公開されている日本語のデータセットを利用して、単語や文章の意味に基づく検索が可能なAIを開発しました。
2. 他機関から公開されている日本語対応のAIと検索性能を比較しました。性能評価には文書データベースから目的文書を検索した際の順位に基づく指標 (MRR) を使用し、キーワード検索において他機関のAI (0.90) に対し開発したAI (0.95) が高い性能を示すことを確認しました。
3. ウェブブラウザによるセマンティック検索の利用を可能とするため、AIを用いた文書検索システムを開発しました。システムをサーバーへ実装し、標準的なCPUで毎秒約7回の問い合わせに回答できることを確認しました。