

生成AIの企業導入で社内情報活用が進む

◆ライオンが「知識伝承のAI化」ツールを自社開発

2023年12月、ライオンは、生成AIと検索システムを用いた「知識伝承のAI化」ツール（図1）を自社開発し、研究開発部門で試験運用を開始した。このツールは、UI（User Interface）、AWS（Amazon Web Services）が提供する生成AIと検索システム、社内の技術文書や実験データなどの情報データベースからなる。利用者はUIへ検索文章を入力するだけで、生成AIが膨大な社内情報から該当文書を取得し、中身を分析・評価して、対話形式で簡潔にまとめた結果を表示する。さらに、関連文書を一通り検索した上での要約や、複数の文書からトレンドを解析して可視化することも可能である。これまで蓄積した社内情報と生成AIを連携したツールにより、業務効率の向上が期待できる。

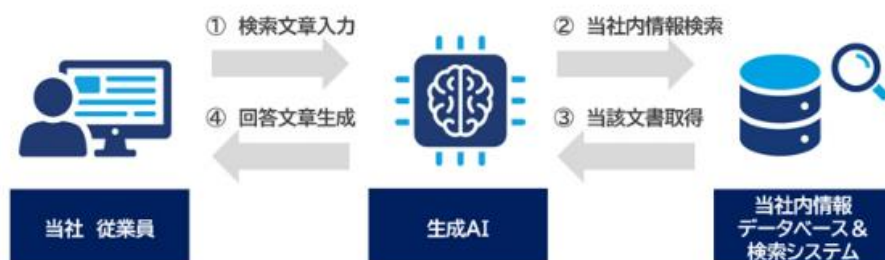


図1 生成AIと検索システムを用いた「知識伝承のAI化」ツール（出典：ライオン）

◆生成AIのビジネス活用が進むと同時に課題も見えてきた

2023年に入り、大規模言語モデル（LLM：Large Language Model）をベースとした生成AIを活用する企業が増えた。生成AIに投げかける指示文（プロンプト）を工夫したプロンプトエンジニアリングによって、文書作成やアイデア出しなどで一定の成果があった一方で、自社固有の情報に関する質問には回答はできない、引用元などが不明で回答の透明性に欠けるといった課題も見えてきた。

そこで、より進んだ生成AIの活用方法を模索する動きが出てきている。生成AIのカスタマイズには複数の類型があり、表1の右に行くほど回答精度は高まるが、開発コスト・難易度とのトレードオフ関係がある。RAG（Retrieval Augmented Generation、検索拡張生成）は、生成AIがプロンプトに答える前に、外部の組織固有の知識データから検索した関連情報をプロンプトと組み合わせる

ことで、一般的な回答ではなく、組織固有の情報に基づく専門的で正確な回答が期待できる。また、知識データを追加学習させるファインチューニングと異なり、生成AIモデルに手を加えないため実装しやすいことも特長だ。

表1 生成AIのカスタマイズ類型（各種資料よりARC作成）

	プロンプトエンジニアリング	RAG	ファインチューニング
データ要件	なし	組織固有の知識データをベクトル化したもの	追加トレーニング用ラベル付きデータセット
トレーニング時間	なし	小～中（データのベクトル化）	中～長（データサイズによる）
開発コスト	低	中	高
回答精度	低	中	高
メンテナンス要件	継続的なプロンプト調整	組織固有の知識データの定期的な更新	定期的なデータ更新と再トレーニング

◆RAG技術で社内情報活用を進めた企業事例

23年10月、セガサミーは、MicrosoftのChatGPTによるチャットボットを機能拡張し、社内情報を参照可能な「IntraSearchモード」を開発した。検索精度を高めることで社内情報の効率的な収集・利用を可能にしており、特に情報のベクトル化とセマンティック検索という2つの技術が重要である。まず、情報のベクトル化では、AIモデルであるGPT-4/Embeddingが社内文書を要約し、数値の配列であるベクトルに変換する。次に、ベクトル化された情報を利用するセマンティック検索で、検索文章および社内文書の単語を文脈から理解して類似性の高い情報を見つける。これらの技術により、単純なキーワードマッチングによる従来の検索方法に比べて、より精度の高い関連情報を検索できる。

23年10月のセミナーで、竹中工務店は、建設業界の熟練者のノウハウ継承と人材育成の課題を解決するための取り組みを発表した。同社は、RAG技術を基に、豊富な社内情報に加えて技術標準書類や過去の建築工事例などのデータに精通し、専門的な指摘を行える「デジタル棟梁」と呼ばれるAIシステムを開発した。

◆ビッグテック各社の提供サービスが充実し、企業の生成AI活用は深化の段階に

今後は、組織固有のデータベースの整備によってRAGの活用が進展すると同時に、GPT-4やPaLM2以外の技術詳細が公開されているオープンなLLMをファインチューニングして用いるなどの動きが活発化すると思われる。AWS、Microsoftなどビッグテック各社から様々なLLMを使いこなすための機能・サービスが提供されており、オープンなLLMも活用の選択肢になり始めている。 【永田紘基】