

AI を活用した建築施工技術探索システム

ます だ きよし
増田 潔

大成建設(株) 技術センター 都市基盤技術研究部長

1 はじめに

建築業界は現在、担い手不足や就労者の高齢化、働き方改革といった環境下で、専門技術をいかに次世代に継承していくかが大きな課題になっている。また、若手社員でなくとも、専門分野に特化した技術情報の著しい増加に伴い、専門技術が属人化して知識共有が困難な状況に陥っている。

このような技術ナレッジ検索の問題を解決する技術として注目されているのが、質問に対して膨大な学習データを基に人間のように解答できる大規模言語モデル(Large Language Model: LLM)、いわゆる生成 AI である。しかし、現状の生成 AI には構造上の問題ともいえるハルシネーション(もっともらしい嘘をつくような現象)があり、現在提供されている生成 AI のサービスを専門技術に対するナレッジ検索に直接使用することは難しい。

このハルシネーションを解決する手段の一つとして、RAG(Retrieval Augmented Generation: 検索拡張生成)が普及しつつある。RAG は膨大な専門知識を効率的に活用し、正確で文脈に応じた情報を提供することができる。大成建設では2022年11月に ChatGPT が公開された直後より、RAG による「建築施工技術探索システム」の開発を内製により開始し、一部社員による試用と改良を繰り返し、2024年7月より建設現場の社員を中心に全社での本格運用を開始した。本稿では、開発したシステムの概要とその運用について報告する。

2 建築施工技術探索システムの概要

開発した施工技術探索システムの概要図を図1に示す。

アプリ上で利用者が質問を入力すると(①)、AI 的検索システムが自社の施工基準・指針・施工事例・技術報告書等が格納されている専門技術データベースから質問に対する関連書類を検索しアプリに送る(②)。アプリケーションは生成 AI に質問と関連書類を入力として与え(③)、これらの情報を基に適切な解答を生成させる(④)。生成された解答と、解答の生成に使用された関連書類をアプリが利用者に提示する(⑤)。ここまでの一般的な RAG と呼ばれる機能である。RAG は、情報検索技術と生成 AI とを組み合わせることで、生成 AI のハルシネーションを低減することが可能な革新的な AI システムである¹⁾。

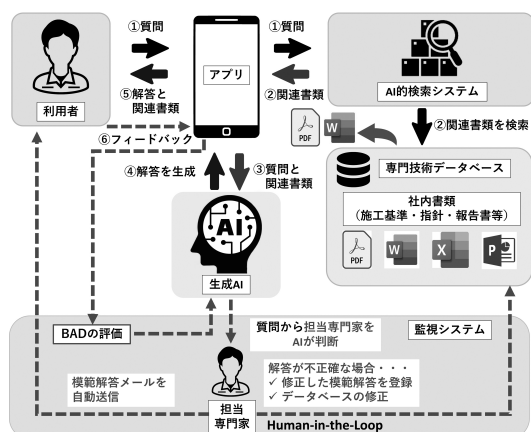


図1 システム概要

しかし、残念ながら RAG の機能のみでは、建設現場で必要とされる解答精度を得られないことが一部社員による試用によって分かったため、本システムには Human-in-the-Loop (AI システムに人間の判断や操作を組み込む思想) に基づく監視システムが追加されている。

監視システムは、大成建設の技術分野を統括する部署の技術者とシステム管理者のみがアクセスでき、RAG によって適切な解答が行われているかチェックすることができる。解答が不正確だった場合や解答に対して利用者が「BAD の評価」をフィードバックした場合に、生成 AI が選択した技術専門家に模範解答作成を依頼する。担当専門家によって作成された模範解答は専門技術データベースに、優先的に検索される模範解答データとして追加されるとともに、質問をした利用者にもメールにて自動送信される。また、データベースの内容そのものに改善点があると判断されれば、担当専門家はデータベースの更新を実施する。

3 開発から運用まで

本システムの開発当初は RAG システムを構築することに注力した。開発には Amazon Web Services (AWS) のクラウドサービスを活用し、セキュアで拡張性の高いアーキテクチャを実現した。

RAG の検索コンポーネントとしては、Amazon Kendra²⁾を採用した。Kendra は機械学習を活用した高度な検索機能を提供し、非構造化データに対しても効果的な検索が可能である。本システムでは検索システムの特徴を最大限に活用するため、収集した社内文書を検索システムのインデックスに登録し、全文検索を可能とした。さらに、文書タイプやデータソース等のメタデータを各文書に付与し、詳細な検索条件の設定を可能にした。そして、文書の機密性に応じたアクセス制御を実

装することで、セキュリティを確保した。

同時に、建設業界の専門用語や略語を含むカスタム辞書も作成し、検索システムに実装した。これにより、「PC 工法」と「プレキャストコンクリート工法」等、建設業界で頻繁に使用される同義語や類語を含む質問に対しても適切な検索結果を提供することが可能となった。

生成 AI モデルについては、Amazon Bedrock が提供する生成 AI のうち、Anthropic 社が開発した Claude シリーズを採用し、常に最新のモデルを使用するようにしている(原稿執筆時には Claude 3.5 Sonnet)。AI モデルのパラメータ調整を綿密に行い、建設業界特有の文脈に適した出力を得られるよう最適化を図った。また、効果的なプロンプトテンプレートを設計し、検索結果と質問を適切に組み合わせる方法を開発した。この過程では、建設業界特有の文脈や要求を十分に考慮したプロンプトを策定した。検索結果の処理においては、検索システムからの出力を生成 AI に入力する前に、関連性や重要度に基づいたフィルタリングと順位づけを実施した。さらに、検索結果の要約を生成する際に入力の処理を最適化することで、生成 AI の性能を最大限に引き出す工夫を行った。

以上のような構成で RAG を構築し、開発部門の従業員、本システムのユーザーとなる部門の幹部、建設現場の従業員から合計60名に限って試験運用を行った。対象データは社内の技術文書約10万ドキュメントとした。

試験運用時には、利用者から「質問に対して的確な解答をしてくれる場合もあるが、的外れな解答もある」という指摘があった。これについて精査した結果、RAG の構造的問題というよりは、データベース側に問題があることが明らかとなった。当初、AI システムということで、大量にデ

ータを与えた方がよい解答を得られるのではないかと考え、必要と思われる技術文書を片端からデータベースに格納したが、その結果、ある質問に対して、あまり役に立たない情報が的確な情報を上回る情報量になり、優良な情報を含んだ技術文書ではなく、内容が劣った技術文書が検索されてしまうことが分かった。そこで、技術文書を精査し、文書間で重複内容がある場合は、より洗練された内容の文書を残すようにして4万5,000ドキュメントまで絞り込んだ。また、ある程度、解答に関係しそうな技術文書の分類が想定される場合には、その分類に限って検索をかけた方が、解答精度が向上するため、検索する文書の種類をあらかじめ絞り込めるフィルタリング機能も持たせた。

RAG システムの改善により、解答精度は明ら

かに向上したが、それでも不正確な解答をしてしまうケースも見られた。最終的には単に RAG システムの改善を行うだけでは、実用に耐え得るシステムとしての維持は困難と判断し、図1に示したように Human-in-the-Loop の考え方を取り入れて監視システムを加え、運用しながらも精度向上を図れるようにした。

これらの機能を持たせたアプリのユーザーインターフェイスを図2に示す。質問は一般の生成 AI と同様に、文章で入力することができる。質問を入力すると数秒で「AI の解答」が表示され、最後に AI に与えられた関連書類の一覧が表示される。「AI の解答」の各文末には解答の生成に利用した関連書類の引用番号が表示されている。この番号は、関連書類の一覧の行頭にある番号と対応しており、どちらかをクリックすれば、その書類の中でも質問の解答根拠となったページを表示する。技術書類の中には数百ページに及ぶものもあるため、単に技術書類を示すだけでなく、知りたい情報があるページをダイレクトに表示するようにした。

以上のような改良をアジャイル的に実施して、2024年7月より全国の建設現場の社員を中心に全社に本システムを公開した。公開から10ヵ月程度経過した現在においても高い利用頻度を保ち続けている。

利用者の声を一部紹介すると、「知りたい知識、情報をスピーディーに教えてくれる本システムは特に若手社員に有用である。ベテラン社員にとっても、客観的かつ最新の情報が確認できてよい」、「膨大な資料を1ページずつ探す手間もなく「〇〇について教えて」と記すだけで、確度の高い解答と根拠が提示され、安心感がある」など、概ね好評である。また、図2にある評価ボタンの「Bad」のフィードバックが徐々に減少し、一方「Good」

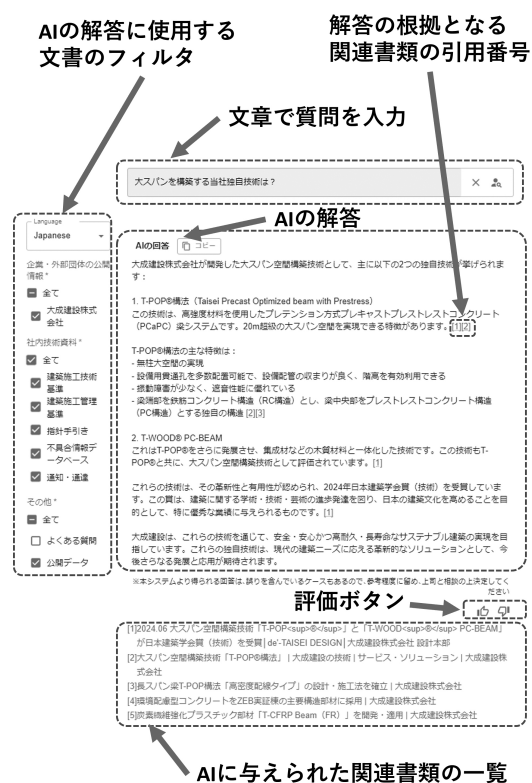


図2 アプリの表示画面と機能

のフィードバックが増えていることが、監視システムがうまく機能していることを裏づけている。

全社公開した後も、内製開発システムである強

みを活かし、常にユーザーからの声を吸い上げて、機動的に改良を継続している。

4 使用例

同じ質問を生成 AI にした場合と本システムにした場合の比較を図 3 に示す。専門的な技術情報の有無を際立たせるために、社内情報に特化した質問をした例である。質問にある「T-iDigital® Field」とは、大成建設の土木現場支援ツールとしてデジタル技術を駆使して開発されたシステムの総称である³⁾。通常の生成 AI では、質問中のシステム名称に似た名称を持つ全く別のシステムについて回答をしているようで、その内容は完全な間違いである。これに対し、本システムでは社内データの基づいた正確な回答をしている。

5 おわりに

生成 AI の進歩は目覚ましいものがあるが、その仕組み自体に、確率的な手法が含まれているため、確度 100% の情報を引き出すことは難しい。文書の要約や、厳密性を求めない情報の生成は公開されている生成 AI で問題なくできるが、特に厳密性を要求される専門技術に対するナレッジ検索については、ここで紹介した建築施工技術探索システムで使用している RAG や Human-in-the-Loop に基づく技術を応用していくことが重要である。

(参考文献)

- 1) Patrick Lewis, Retrieval-Augmented Generation for Knowledge-Intensive, NLP Tasks (2020)
<https://arxiv.org/abs/2005.11401>
- 2) Amazon Kendra
<https://aws.amazon.com/jp/kendra/>
- 3) 「T-iDigital®Field」の施工管理および安全管理機能を拡張、大成建設ウェブサイトニュース (2022. 4. 26)
https://www.taisei.co.jp/about_us/wn/2022/220426_8777.html

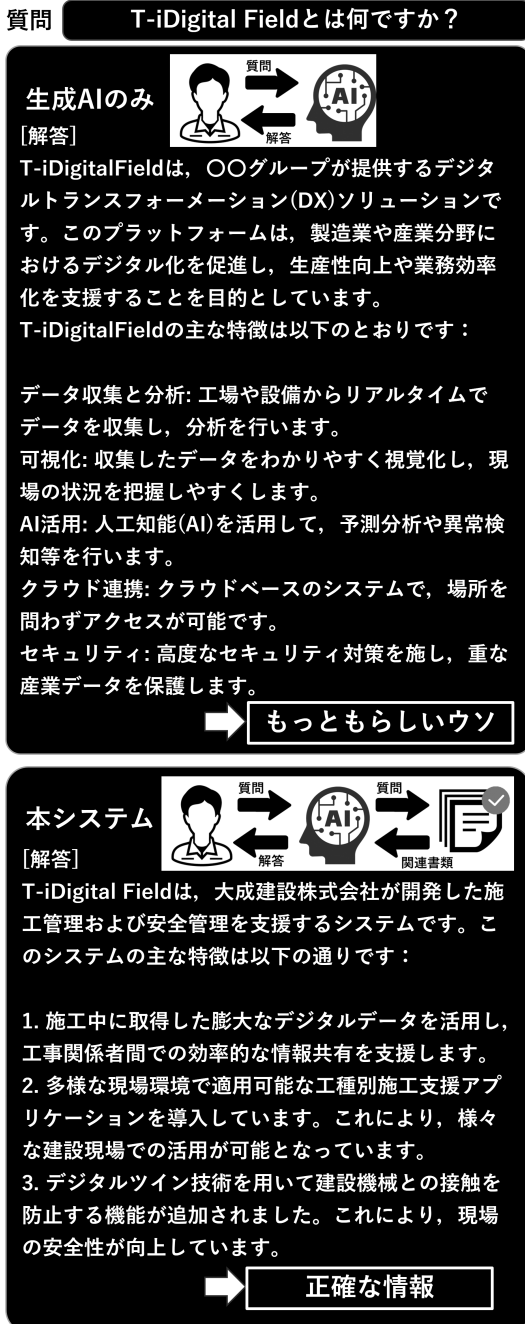


図 3 本システムと生成 AI のみの解答の比較