2024 年 5 月 16 日
株式会社ヘッドウォータース
(コード番号：4011 東証グロース)

生成 AI の業務活用における「RAG がうまくいかない」企業向けに Microsoft Fabric をデータプラットフォームとした「Advanced RAG」サービス開始

AI ソリューション事業を手掛ける株式会社ヘッドウォータース（本社：東京都新宿区、代表取締役：篠田 庸介、以下「ヘッドウォータース」)は、生成 AI の業務活用における「RAG がうまくいかない」企業向けに、日本マイクロソフト株式会社（本社：東京都港区、以下「日本マイクロソフト」）が提供する「Microsoft Fabric」をデータプラットフォームとした「Advanced RAG」サービスを開始したことをお知らせいたします。



ヘッドウォータースでは、「Azure OpenAI Service」による企業向け GPT サービスラインナップ拡充を行い、企業向けに生成 AI ならびに LLM（大規模言語モデル）と当社の技術力を活かした RAG(Retrieval Augmented Generation)システムや伴走支援型ラボを多くの企業向けに提供してまいりました。

RAG を利用する多くのお客様から、課題として「RAG がうまくいかない」という相談を多く受ける事となり、当社ではこの課題に対してソリューションを模索してまいりました。

RAG（企業独自カスタムの生成 AI）において学習データの元となるのが Microsoft Word や Microsoft PowerPoint、PDF と言った非構造化の社内ドキュメントですが、読み込むドキュメントが多くなると「学習しなくてよい情報まで学習してしまう」事でノイズが発生することが多くあります。

RAG がうまくいかない理由のうち、大きな要因として、生成 AI に学習させる社内ドキュメント類が整備されていないことにあります。

- 社内ファイルを組織横断した場合、命名規則や用語が部署やプロジェクトごとに違う
- 社内向けに作られたファイルなのでレビューが通っておらず、誤字脱字も含まれている
- 社内ファイルのバージョン管理までされておらず、最新版ではない過去のファイルが残っている
- 社内ファイルにおける優先順位付けや重みづけまで定まっていない。
- Word や PowerPoint ファイルに含まれる画像オブジェクトを説明する文章が無い。

社内ファイルはそれぞれ個別用途で作成しており、個人が自由度高く用意できるメリットがあるものの、データエンジニアリングの世界からするとデメリットにもなります。

一方、しっかりとしたデータ整備やルール化を行うには大掛かりになりすぎて現実的ではなく、曖昧な所を汲み取って欲しいというニーズが生成 AI に期待されている領域でもあります。

このような課題解決として、ヘッドウォータースは Microsoft Fabric をデータプラットフォームとした「Advanced RAG」サービスを提供することとなりました。

本サービスの特徴は、直接的に社内ファイルを生成 AI に読み込ませるのではなく、RAG の精度を上げる為に「キレイなデータ」と「ファイルメタデータ」を蓄積するデータプラットフォームを用意し、RAG の前工程と後工程におけるデータ処理を行います。

Microsoft Office 365 ファイルには「ファイル作成者」「更新者」「更新日付」「アクセス日付」「バージョン」と言ったプロパティ情報があり、社内ファイル共有として多くの企業が利用している Microsoft SharePoint では「閲覧回数」「平均使用時間」「人気度」「社内シェアされたか否か」といったアクティビティデータが残されています。

これらのメタデータを取ることで、社内ファイルの権威性や信頼性、人気度から重みづけスコアを発行し、スコア上位に絞り込んだ上で生成 AI に学習させることで、RAG の精度を向上させます。

また、生成 AI を利用して予め要約を施し、要約情報から検索を行う「Summary Embedding」を行うことで、余計なデータを取り込まずに済み、精度向上だけでなくレスポンスタイムの向上やクラウドコストの削減にも寄与することができます。

現在多くの企業で利用している Microsoft Office 365 ファイルを簡単に取り込めるデータプラットフォームは、日本マイクロソフトが提供する「Microsoft Fabric」であり、様々な「Advanced RAG」を必要とする生成 AI シナリオに対応します。

「Advanced RAG」サービスは、すでに複数の大手企業から受注が決定しております。

■エンドースメント

日本マイクロソフト株式会社
執行役員 常務 クラウド＆ソリューション事業本部長 岡嵜 禎様のコメント

この度は、Azure OpenAI Service と Microsoft Fabric を基盤としたソリューションをご提供頂けることを心から御礼を申し上げます。生成 AI の導入効果を出していくには、参照先の構造化データ・非構造化データの整備が不可欠です。ヘッドウォータースが提供する「Advanced RAG」は、Microsoft Office に溜まっている社内業務ノウハウを活用したいという利用シナリオで、多くのお客様が課題と感じている RAG 精度向上に対して、再現性がある解決策を提示してくれています。弊社製品とヘッドウォータースの独創性・技術力と組み合わせることで、今後ますます企業変革を加速できると確信しております。

■Advanced RAG とは

高度な検索拡張生成（ARA）は、より洗練されたデータ処理を使用して検索品質を向上させる生成 AI 技術です。主に「検索読み取り」プロセスを伴う「Naive RAG」の問題に対処します。ARA は複数または反復的な検索を使用し、次のような技術が含まれます。

- 検索前の最適化（Pre-Retrieval）:文ウィンドウ検索など
- 検索の最適化(Retrieval):ハイブリッド検索など
- 検索後の最適化(Post-Retrieval):再ランク付けなど
- インタラクティブ RAG:ユーザーがリアルタイムで検索プロセスに積極的に影響を与えることができる
- 取得したデータの統合強化:多くの場合、コンテキストと関連性の重み付けを伴う
- 反復的な改良：精度と関連性の向上が可能

■Microsoft Fabric とは

Microsoft Fabric は、データ移動からデータ サイエンス、データ レイク、データ エンジニアリング、データ統合、リアルタイム分析、ビジネスインテリジェンスまですべてをカバーする企業向けオールインワン分析ソリューションの SaaS サービスです。

Microsoft Fabric は、複数ベンダーのさまざまなサービスをつなぎ合わせる必要がなく、分析ニーズを簡素化し、生成 AI を中心とした AI 連携に優れています。

Copilot in Fabric では、自然言語を中心としたデータ分析を可能としています。

https://www.headwaters.co.jp/news/prompt_engineering_labo.html

■RAG(Retrieval Augmented Generation)とは

Retrieval Augmented Generation（RAG）は、大規模言語モデル（LLM）と外部のデータベースや情報源を結びつけるための新しい技術です。外部の知識ソースを検索し、より強化した文章生成を行います。

■Azure OpenAI Service とは

OpenAI 社とマイクロソフト社のパートナーシップから生まれた「Azure OpenAI Service」は、さまざまなユースケースに応じて特定のニーズに合わせた最適化が可能となる業界をリードするコーディングおよび言語

AI モデルを提供するサービスです。セキュリティー、信頼性、データプライバシー、ID 認証、有害コンテンツのフィルタなどエンタープライズ企業が OpenAI Service を利用するための Azure マネージドサービスです。

https://azure.microsoft.com/ja-jp/products/cognitive-services/openai-service

■参考情報

Azure OpenAI Service リファレンスアーキテクチャの Advanced パートナー認定について

https://www.headwaters.co.jp/news/azure_openai_service_advanced_partner.html

■今後について

Microsoft Fabric を活用した「Advanced RAG」は企業独自の生成 AI 活用におけるすべての RAG プロジェクトに適用できると考えております。

今後はより多くのデータを処理するための Azure Databricks 連携、データガバナンス強化のための Microsoft Purview 連携、ユーザーデータを加味する Microsoft Entra ID 連携、RAG 前データのバッチ処理などを行っていき、企業内利用に留まらず、BtoC モバイルアプリやスマートグラス、エッジ AI への適用などにソリューション展開を図ってまいります。

また、当社の掲げるアライアンス戦略において顧客企業ともビジネスパートナーとなり、共に生成 AI 経済圏を拡大する取り組みを行ってまいります。

なお、本件による当社の当期業績に与える影響は軽微であります。今後開示すべき事項が発生した場合には速やかにお知らせいたします。

■商標について

Microsoft Azure、PowerPoint、Office 365、SharePoint、Microsoft Purview、Microsoft Entra は、米国 Microsoft Corporation の米国およびその他の国における登録商標または商標です。その他、記載されている製品名などの固有名詞は、各社の商標または登録商標です。

<会社情報>

会社名：株式会社ヘッドウォータース

所在地：〒163-1304　東京都新宿区西新宿 6-5-1 新宿アイランドタワー 4 階

代表者：代表取締役　篠田 庸介

設立　：2005 年 11 月

URL　：https://www.headwaters.co.jp

<本件のお問い合わせ>

株式会社ヘッドウォータース

メール：info@ml.headwaters.co.jp