

株式会社エム二

Azure AI を活用したRAGの精度改善



自己紹介

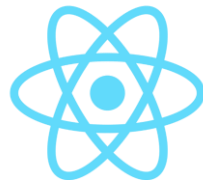


エンジニア

宮木翔太

SHOTA MIYAKI

- 主にRAGを中心にWebアプリのインフラ周りの設計・構築・運用などを主に担当
- 技術スタック



1. Why Azure ?
 - Azure AI Search
2. GraphRAGの活用
 - グラフ構造を利用した検索
 - 強み
 - ユースケース
3. Prompt Flowでの精度改善
 - 評価フロー
 - 強み
 - 実演 : GraphRAGの出力をPrompt Flowで評価してみよう

Why Azure ?

- 検索性能の良さ

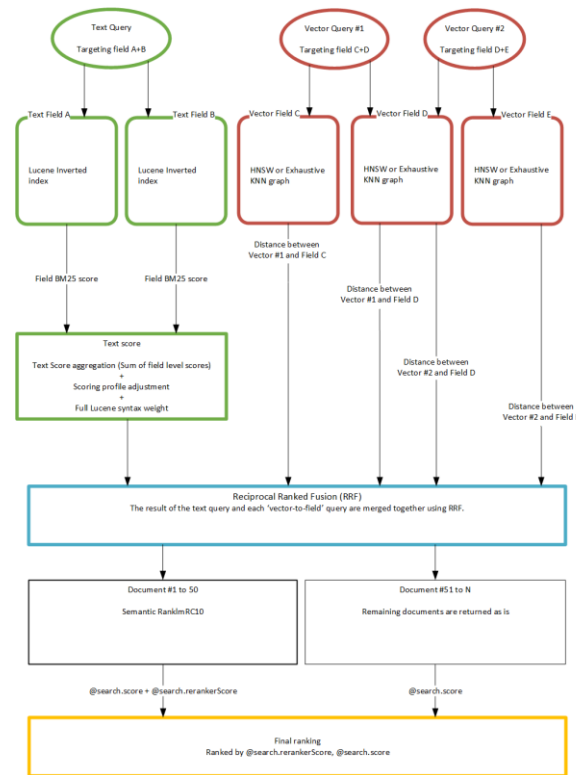
- セマンティック検索
 - Bing検索にも使用されているセマンティックアルゴリズム
- ハイブリッド検索
 - キーワード検索+ベクター検索
 - RRF手法

- ストレージ

- Azure Blob Storage
- SQL
- Data Lake
- etc

- その他AIサービスとの統合によるエンリッチメント

- OCR(Azure Document Intelligence)
- 翻訳
- etc



**ベクトルDB・検索サービスはRAGの中核であり、土台である
土台がしっかりしていないと精度改善はなし得ない**

Graph RAGの活用

グラフ構造を利用した検索

- LLMを用いて膨大な非構造データからエンティティと関係を抽出し、ドキュメント全体のナレッジグラフを自動構築

例：

「日本の首都は東京である」



- 要素：日本，首都，東京
- エンティティ：日本，東京
- リレーション：首都

ナレッジグラフ

- ベクトルDB作成時のチャンク分割によって，切り落とされてしまうチャンク間の関連性を掬い上げることができる

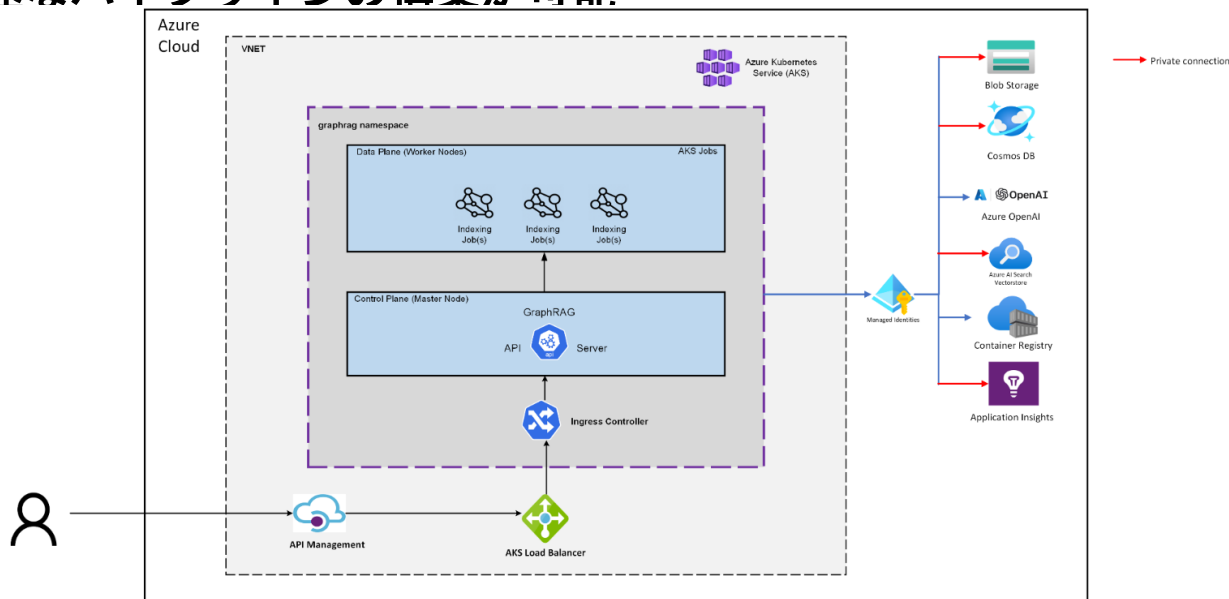
- ベクトル検索中心のアプローチでは見落としがちな関連情報を知識グラフ経由で見つけ出し、LLMに的確な文脈を提供することで、回答の正確さや一貫性を高める
- 例えば. . .
- 社内ドキュメント（設計書、手順書）、ナレッジベース（FAQ、ベテラン社員の報告書）、さらにはTeamsやSharePoint上の議事録まで幅広くデータを繋ぐと、GraphRAGに「生産ラインで発生したある不具合の根本原因と対策」を質問した際に、関連する過去のトラブル事例・修理履歴・専門家の議論を知識グラフ上で辿り包括的な回答根拠を提示が可能

- 以下のような調査はナレッジグラフを発展させたGraphRAGの得意とするところ。
 - 水平展開により関連する事象を発見すること
 - 共通のパターンを認識すること
 - 背景の文脈を理解すること
 - 多角的な視点に気づくこと

「人手不足の領域にて、
初学者が生成AIを活用して専門的なドメイン知識を獲得していく」
といったようなユースケースがフィット

- GraphRAG Accelerator

- GraphRAGの動作に必要なリソース一式をAzure上にデプロイするためのIaCツールが用意されており、手順に従うだけで簡単に試すことが可能
- 迅速なパイプラインの構築が可能



Emniでのユースケース

- 部署を横断した様々な社内申請書の利用目的、申請手順などの情報をGraph RAGによりインデックス化
- 各部署に散らばった申請書情報を1つのナレッジグラフに集約

わずか1ヶ月の開発期間で
事前に想定されていた質問(10問)に対する
回答精度100%を達成！

Prompt Flowを使用した精度改善

- Prompt Flowでは出力の品質を測る評価フローを作成することができ、モデルの回答が期待値をどれだけ満たすか評価指標（メトリクス）を自動で算出
- 様々な評価フローのテンプレートがある

ギャラリーを探す

All Standard flow Chat flow **Evaluation flow**

Evaluation Classification Accuracy Evaluation

Measuring the performance of a classification system by comparing its outputs to groundtruth.

詳細の表示 クローン

Evaluation QnA Groundedness Evaluation

Compute the groundedness of the answer for the given question based on the context.

詳細の表示 クローン

Evaluation QnA Relevance Evaluation

Compute the relevance of the answer for the given question based on the context.

詳細の表示 クローン

Evaluation QnA Coherence Evaluation

Compute the coherence of the answer base on the question using llm.

詳細の表示 クローン

Evaluation QnA Fluency Evaluation

Compute the Fluency of the answer base on the question using llm.

詳細の表示 クローン

Evaluation QnA Ada Similarity Evaluation

Compute the cosine similarity between the answer and the ground truth embedded with ada embedding.

詳細の表示 クローン

Evaluation QnA GPT Similarity Evaluation

Compute the similarity of the answer base on the question and ground truth using llm.

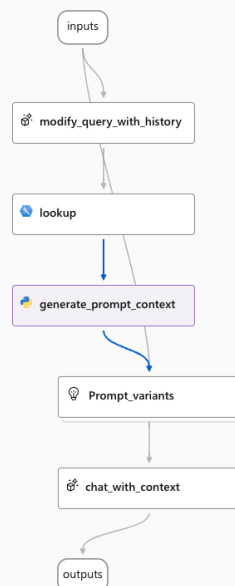
詳細の表示 クローン

Evaluation QnA F1 Score Evaluation

Compute the F1 Score based on words in answer and ground truth.

詳細の表示 クローン

グラフ



- 一連の評価プロセスを「フロー」として可視化・管理できる
 - 整理しやすく，再現性が高い
- どの変更が有効かを定量的に比較検証でき、最適な設定をすばやく見極められる
 - この変更をすると質問Aに答えられるようになったけど，質問Bには答えられなくなった，など
- CLIでの実行もできるのでCI/CDに組み込める

実演：GraphRAGの出力をPrompt Flowで評価してみよう

● 用意したRAG

- データセット：wikiからスクレイピングした「生産方式」「品質管理」などに関連する記事17件
- GraphRAG AcceleratorでデプロイしたAzure上のGraphRAG

インダストリー4.0

この項目「インダストリー4.0」は翻訳されたばかりのものです。不自然あるいは曖昧な表現などが含まれる可能性があります。このままでは読みづらいかもしれません。(原文: en:Industry 4.0) 6/52, 18 April 2019)

修正、加筆に協力し、現在の表現をより自然な表現にして下さる方を求めています。ノートページや履歴も参照してください。(2019/4/4)

インダストリー4.0 (英: Industry 4.0、略称: I4.0) とは、製造業におけるオートメーション化およびデータ化・コンピュータ化を目指す昨今の技術的コンセプトに付けられた名称である。具体的には、サイバー・フィジカルシステム (CPS)、モノのインターネット (IoT)、クラウドコンピューティング[1][2][3][4]、コグニティブコンピューティングなどが含まれる。インダストリー4.0は一般に第四次産業革命として言及される[5]。



サプライチェーン・マネジメント

出典は参照するだけでなく、脚注などを用いてどの記述の情報源であるかを明記してください。記事の信頼性向上にご協力をお願いいたします。(2018/12/4)

サプライチェーン・マネジメント (英: supply chain management, SCM)、**供給連鎖管理** (きょうきょくせんさかんり) とは、商品の原材料調達から生産加工や在庫管理、流通や販売、各プロセスに携わる物流など、商品の開発から消費者の手に渡るまでの一連の流れを指す[1]。また、複数の企業間で統合的な物流システムを構築し、経営の成果を高めるためのマネジメント手法ともなる。なお、この場合の「複数の企業」とは同業の親会社・子会社のような企業グループ内での関係に限らず、対等な企業間で構築される物流システムもサプライチェーン・マネジメントと呼ばれる[2]。しかし、実務上は企業間の取引は対等とせられず、販売と理論との乖離があり、その関係[3][4]分析が重要になる。また、サプライチェーンが顧客に届くまでの複数企業間の流れであるのに対して、バリューチェーンは一企業内の業務の流れを指す。

米国のサプライチェーンカンファレンス (SCML) による定義では「後援提供活動の初めから終わりまで、つまり原材料の供給者から最終消費者に至る全過程の個々の業務プロセスを、一つのビジネスプロセスとしてとらえ直し、企業や組織の壁を超えてプロセスの全体最適化を継続的に行い、製品・サービスの顧客付加価値を高め、企業に高収益をもたらす戦略的な経営管理手法」とされている[5]。

ntのエンドポイント経由で

フォードイズム

この項目「フォードイズム」は翻訳されたばかりのものです。不自然あるいは曖昧な表現などが含まれる可能性があります。このままでは読みづらいかもしれません。(原文: en:Fordism) 4/16, 18 April 2019)

修正、加筆に協力し、現在の表現をより自然な表現にして下さる方を求めています。ノートページや履歴も参照してください。(2019/4/4)

現代の資本主義を特徴付ける概念であり、経済学のレギュラシオン理論や、社会学、さらに経済地理学などで言及されることが多い。もともとは、イタリアの思想家アントニオ・グラムシの命名によるとされている。

トヨタ生産方式

出典は参照するだけでなく、脚注などを用いてどの記述の情報源であるかを明記してください。記事の信頼性向上にご協力ください。(このテンプレートの使い方)

出典欄へ「トヨタ生産方式」- ニュース、書籍、スカラー- JSTOR、NDL、dlib.jp、ジャーナルネット、TDBL (2013年8月)

トヨタ生産方式 (トヨタせいさんほうしき、Toyota Production System、略称TPS) は、トヨタ自動車の生み出した、工場における生産活動の運用方式の一つ。多くの企業がこれにならった方式を取り入れており、工場等の製造現場やそれに付随するスタッフ部門だけでなく、間接部門でも取り入れている企業も見られる。

河合隆によると、TPSが普及する以前も「必要の時に必要なだけ作る」という考えが現場に浸透していたという[1]。

rg-graphrag-webinar リソースグループ

検索

作成 ビューの管理 リソース グループの削除 更新 CSV にエクスポート クエリ

基本

サブスクリプション (移動): [subscription-emu2-main](#)

サブスクリプション ID : 8c88c38f-cbd7-4efd-a667-76615651ee21

タグ (編集) : [タグの追加](#)

リソース 推奨事項

任意のフィールドのフ... 種類 次の値と等しい すべて × 場所 次の値と等しい すべて × すべて フィ

33 件中 1 ~ 33 件のレコードを表示しています。 ☐ 非表示の型の表示 ☐

- ☐ 名前 ↑
- ☐ aks-buShcpcpkxew
- ☐ aks-publickey
- ☐ aoi-graphrag-webinar
- ☐ apim-buShcpcpkxew
- ☐ appi-buShcpcpkxew
- ☐ Containerinsightslog-buShcpcpkxew
- ☐ cosmos-buShcpcpkxew
- ☐ crbuShcpcpkxew
- ☐ FailureAnomalies_appi-buShcpcpkxew
- ☐ graphrag.io
- ☐ id-buShcpcpkxew
- ☐ log-buShcpcpkxew
- ☐ nsg-buShcpcpkxew

実演：GraphRAGの出力をPrompt Flowで評価してみよう

● 出力を見てみる

```
%%time
# pass in a single index name as a string or to query across multiple indexes, set index_name=[myindex1, myindex2]
local_response = local_search(
    index_name=index_name,
    query="フォードイズムとは何か?",
)
```

```
print(local_response.json()["result"])
```

✓ 18.1s

Python

フォードイズムとは

フォードイズム (Fordism) は、アメリカの実業家ヘンリー・フォードが自動車工場で考案した生産方式および経営思想を指します。この概念は、現代資本主義を特徴付ける重要なモデルとして、経済学や社会学で広く言及されています [Data: Entities (1177, 642); Relationships (615, 630)]。

フォードイズムの中心的な特徴は、大量生産と効率化です。具体的には、科学的管理法を応用し、ベルトコンベアを中心とした生産システムを導入することで、製品の単純化や部品の標準化を実現しました。この仕組みにより、生産率が向上し、労働者の賃金も生産性に比例して上昇する「生産性インデックス賃金」が採用されました。この賃金制度は労働者の士気を高め、購買力を

経済的・社会的影響

フォードイズムは、第二次世界大戦後から1970年代までの高度経済成長期において、経済体制としても重要な役割を果たしました。この時期には、大量生産と消費拡大の好循環を生み出すために、労働組合の承認、最低賃金制度の確立、ケインズ経済政策や社会保障政策が導入されました。これらの要素が組み合わさり、内包的な量産体制と独占的な調整様式を特徴とする経済モデルが

フォードイズムの影響と後継モデル

フォードイズムは、後の生産方式にも影響を与えました。例えば、トヨタ自動車の生産方式 (トヨタイズム) はフォードイズムの影響を受けたとされています。また、フォードイズムの後には、柔軟性や多様性を重視する「ポスト・フォードイズム」という新たな生産方式が登場しました [Data: Entities (649, 657); Relationships (628, 629)]。

さらに、フォードイズムは文化的な影響も及ぼしました。例えば、SF小説『すばらしい新世界』では、フォードイズムに支配された社会が描かれています。この作品は、フォードイズムの社会的影響を考察する上で重要な資料とされています [Data: Entities (656); Relationships (628)]。

結論

フォードイズムは、単なる生産方式にとどまらず、経済、社会、文化に広範な影響を与えた重要な概念です。その効率的な生産システムは、現代の資本主義の基盤を築き、後の生産方式や経済モデルにも影響を与えました。ヘンリー・フォードの革新は、20世紀の産業と社会の発展に大きく寄与したと言えるでしょう。

CPU times: user 8.39 ms, sys: 10.1 ms, total: 18.5 ms

Wall time: 18.1 s

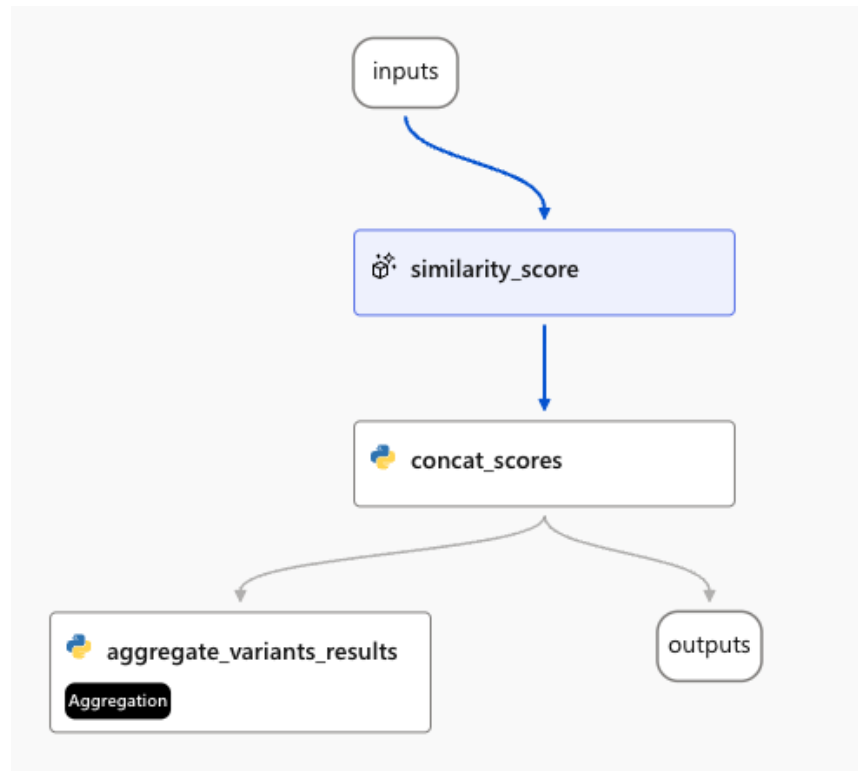
実演：GraphRAGの出力をPrompt Flowで評価してみよう

- テストデータセット20問
 - o1モデルDeep Researchで作成後，ファクトチェック
 - 実際のプロジェクトではデータセットに関する専門家の元で作成

"question"	"answer"
"フォーディズムとは何か？"	"フォーディズムとは、ヘンリー・フォードが自動車工場で実践した大量生産の手法およびそれに基づく経営思想のことです。ベルトコンベアによる流れ作業や部品の標準化を特徴とし、現代資本主義の生産モデルの一つとして知られています。"
"フォーディズムと科学的管理法の関係は何か？"	"フォーディズムは、テイラーの科学的管理法の原理を応用して確立された生産システムです。ヘンリー・フォードは科学的な作業分析と管理手法を組み合わせることで、自動車の大量生産を効率化し、フォーディズムという生産方式を確立しました。"
"科学的管理法とは何か？"	"科学的管理法とは、フレデリック・テイラーが20世紀初頭に提唱した、作業を科学的に分析して効率化するための管理手法です。作業の標準化や時間研究など客観的な基準でムダを排除し、生産性向上と労働者の賃金上昇を追求しました。"
"ジャストインタイム生産方式とは何か？"	"ジャストインタイム生産方式（JIT）とは、生産工程において必要なものを必要な時に必要な量だけ生産・供給することで、在庫のムダを徹底的に減らす生産管理方式です。トヨタ自動車が発見した方式で、「必要なものを必要な時に必要な量だけ生産・供給する」というのが基本理念です。"
"TQM（総合的品質管理）とは何か？"	"TQM（総合的品質管理）とは、全社的に品質向上に取り組む経営手法で、組織全体で統一した品質目標を経営戦略に組み込むものです。すべての部門と従業員が継続的な改善（カイゼン）と顧客満足の向上を追求することを目指します。"
"TQMとシックス・シグマの違いは何か？"	"TQM（総合的品質管理）が企業全体の文化として継続的な品質改善を目指す包括的な管理哲学であるのに対し、シックス・シグマは統計手法に基づいて工程のばらつきを減らし、不良を極限まで低減することに特化したアプローチです。"
"シックス・シグマとは何か？"	"シックス・シグマとは、1980年代に米国のモトローラ社が開発した品質管理手法で、工程の不良発生率を100万回に3.4回以下にまで減らすことを目標としています。統計的なデータ分析に基づいて工程のばらつきを抑え、品質向上を図ります。"
"QCサークルとは何か？"	"QCサークル（品質管理サークル）とは、職場内の小グループが自主的に集まり、品質向上のための問題解決や業務改善に取り組む活動のことです。全員参加のボトムアップ型活動であり、QC手法を用いて現場の作業を継続的に改善します。"
"ロジスティクスとサプライチェーン・マネジメントはどう異なるか？"	"ロジスティクスは、原材料の調達から製品の配送・販売までの物流（物の流れ）を計画・管理することで、モノを効率よく届けることに重点を置いた概念です。一方、サプライチェーン・マネジメント（SCM）は、物流だけでなく、情報流や金銭流の統合的な管理を指します。"
"工作機械とは何か？"	"工作機械とは、金属や木材などの素材に切削・穴あけ・研削などの加工を施して目的の形状に仕上げるための機械装置です。旋盤やフライス盤、ボール盤など様々な種類があり、一般に工具や加工物を動かして材料を削り取り、加工を行います。"
"産業用ロボットとは何か？"	"産業用ロボットとは、人間の代わりに工場などでの作業（例：溶接、塗装、組立、搬送）を自動で行う多関節型の機械装置です。プログラム可能なマニピュレータ（ロボットアーム）を備え、生産ラインの自動化や効率化に大きく貢献しています。"
"3Dプリンターはどのような用途に利用されていますか？"	"3Dプリンターは、現在主に製品開発時の試作品（プロトタイプ）製作に活用されています。また、近年では従来の加工が難しい複雑形状部品の少量生産にも用いられ、自動車・航空宇宙分野の部品製造や医療分野での人工臓器の作成など、幅広い分野で利用されています。"
"産業革命とは何か？"	"産業革命とは、18世紀後半から19世紀にかけて主にイギリスで起こった産業構造の劇的な変革のことです。蒸気機関に代表される石炭エネルギーの活用と機械の導入によって生産力が飛躍的に高まり、社会経済の体制が農業社会から工業社会へと大きく変化したことを指します。"
"インダストリー4.0とは何か？"	"インダストリー4.0とは、製造業における高度なオートメーション化とデジタル化を推進するための近年の技術コンセプトで、いわゆる「第4次産業革命」を指します。IoT（モノのインターネット）やAI、クラウドコンピューティングなどのデジタル技術の導入によって、生産工程の自動化・最適化が進められています。"
"改善（カイゼン）とは何か？"	"改善（カイゼン）とは、生産現場や業務プロセスにおいて継続的に効率化や品質向上を図る取り組みの総称です。現場の作業員が自ら知恵を出し、設備の改良や手順の見直しなど小さな工夫を積み重ねていくボトムアップ型の活動です。"
"システム工学とは何か？"	"システム工学とは、複数の要素から構成される複雑なシステムを効果的に設計・管理するための学際的な工学分野です。個々の部品の設計だけでなく、それらを統合してシステム全体を最適化し、要求された機能や性能を実現するための総合的なアプローチです。"
"持続可能性とは何か？"	"持続可能性とは、現在の社会や環境の仕組みを将来にわたって維持続けられる能力や状態を指す概念です。特に環境分野では、生態系がその多様性と生産性を将来にわたって損なわずに保ち続けられることを意味し、経済活動と環境保護の両立を追求します。"
"労働安全衛生とは何か？"	"労働安全衛生とは、働く人の安全・健康・福祉を職場で守ることを目的とした分野です。職場の危険を予防し、快適で健康的な就業環境を確保することで労働者の生命と健康を守る取り組みであり、企業にとって倫理的・法的な義務でもあります。"
"シックス・シグマプロジェクト手法（DMAIC）はどのような手順で進められるか？"	"シックス・シグマでは、DMAICと呼ばれる5つのステップに沿ってプロジェクトを進めます。すなわち、最初に課題を定義（Define）し、現状を測定（Measure）してデータを収集・分析（Analyze）し、対策を実行して改善（Improve）し、成果を定着（Control）させるというプロセスです。"
"産業革命とインダストリー4.0はどのように違うか？"	"産業革命は18～19世紀に蒸気機関などの新技術によって工業化が進んだ歴史上の大変革を指します。一方、インダストリー4.0は21世紀におけるIoTやAIなどの技術導入による製造業の変革で、「第4次産業革命」と称されています。"

- 評価

- QnA GPT Similarity Evaluation
テンプレートを使用
- 実際の出力とテストデータの回答をGPT-4oで比較
- 1から5の類似度スコアを算出



実演：GraphRAGの出力をPrompt Flowで評価してみよう

● 評価結果

- 類似度平均4.75 正答率 95%

出力の視覚化

最新の情報に更新

共有

実行とメトリック

表示名

状態

作成日

トークン数

gpt_similarity

gpt_similarity_pass_rate(%)

Flow-created-on-04-08-2025-19-18-23-variant_0-202504081225-QnA GP1	Completed	Apr 8, 2025 21:25 PM	0	4.75	95
--	-----------	----------------------	---	------	----

エクスポート

出力

エクスポート

すべてのセルを展開する

検索

フィルター

列

#	実行	状態	入力.question	入力.ground_truth	入力.answer	gpt_similarity	システム メトリクス
0	Flow-created-on-04-08-23	Completed	"フォードイズムとは何か?"	"フォードイズムとは、ヘンリー・フォードが自動車工場です。大規模生産の手法および"	["### フォードイズムとは何か?"\n\nフォードイズム (Fordism) は、アメリカの自動車	5	Duration : 1.100 秒 Token : 0
1	Flow-created-on-04-08-23	Completed	"フォードイズムと科学的管理法の関係とは何か?"	"フォードイズムは、テイラーの科学的管理法の原理を応用して確立された生産システムで"	["### フォードイズムと科学的管理法の関係\n\nフォードイズムと科学的管理法は、20世	4	Duration : 0.9864 秒 Token : 0
2	Flow-created-on-04-08-23	Completed	"科学的管理法とは何か?"	"科学的管理法とは、フレデリック・テイラーが20世紀初頭に提唱した、作業を科学的に"	["### 科学的管理法とは\n\n科学的管理法 (英: Scientific Management) は、フ	5	Duration : 1.151 秒 Token : 0

前へ

次へ

継続的, 定量的な比較検証サイクルを実現することで
良質な精度改善を進めることができる

Thank you.

