



GraphRAG MCP

マルチモーダルナレッジグラフエンジン

Neo4j · LangGraph · React · Vite · Qwen · vLLM · llama.cpp

概要

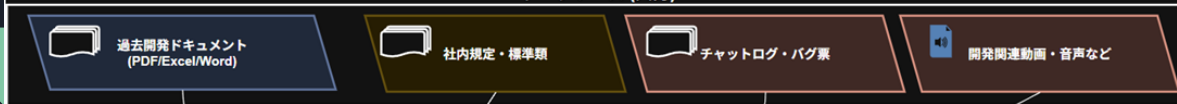
AI技術の進化により、企業規模を問わず高度なシステム構築が可能となり、AI製品の機能も広がり続ける中、ソフトウェア開発における企業間の格差は大きく縮小しています。

そうした中で、機動力は高くともデータを持たない新興企業に対し、歴史ある大手企業が持ち続けている優位性が「膨大な社内データ資産」です。

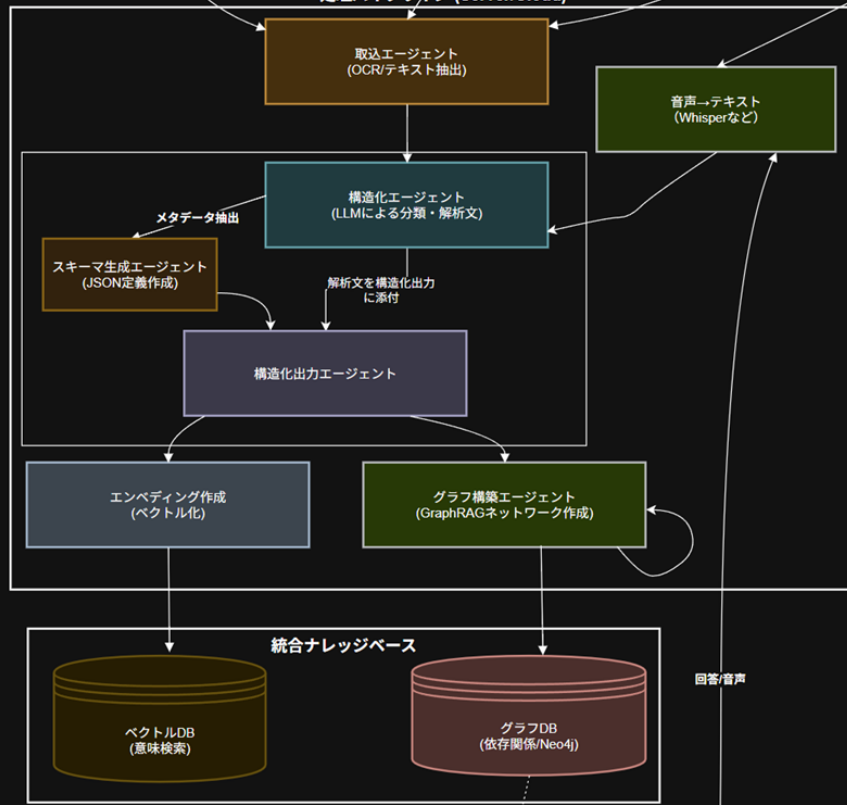
現在の社内ワークフローに合致した質の高いOUTPUTを引き出すためには、このデータ資産をAIへのINPUTとして、安全かつ効果的に活用できるかどうかを鍵を握っています。

データ処理パイプラインと統合ナレッジベースの部分に当たるPoCを開発。

データソース (入力)



処理パイプライン (Server/Cloud)



ドキュメント摂取 & 構造化パイプライン

過去の開発ドキュメントや社内規定を学習し、検索可能な知識ベースを構築するシステムです。

- **目的:** 過去資産のデジタル化、規定検索 (Perplexityライクな根拠提示)、ベテラン知見の保存。
- **エージェント構成:**
 - **取込エージェント:** OCR (Adobe PDF APIやTesseract) を使い、紙/PDFからテキスト抽出。
 - **構造化エージェント:** LLMを使い、非構造化テキストを「規定」「仕様」「図表」等のクラスに分類し、JSONスキーママッピング。
 - **グラフ構築エージェント:** ドキュメント間の参照関係 (「規定Aは規定Bに基づく」など) を抽出し、Neo4j等のグラフDBに格納。
 - **インタビュアーエージェント:** 知識の欠損 (Graph上の孤立ノード) を検知し、ベテランSEに対して質問を生成・実施する。
- **特徴:**
 - **ソーストラッキング:** チャンク (文章の断片) ごとに「ファイル名・ページ番号・条項ID」をメタデータとして付与し、回答時に引用元を明示 (Perplexity機能)。

使用例: MCPサーバーによるナレッジ検索

開発経験2年必須のPython案件にマッセン...

```
</> g:\RAG_mcp\README.md
```

開発経験2年必須のPython案件にマッチングしている技術者を見つけ、テーブルでまとめてください

● Rag-local [get_document_graph]

```
OUT {"document_id": "72b23cc1-a65d-4037-87c4-e308bf4c6714", "nodes": [{"id": "72b23cc1-a65d-4037-87c4-e308bf4c6714", "label": "skillsheet_40_Data_Engineer_Spark.md", "type": "Document", "title": "skillsheet_40_Data_Engineer_Spark.md"}]}
```

● Thought for 15s >

● RAGシステムから抽出したデータに基づき、Python開発経験2年以上の技術者をまとめます。

Python開発経験2年以上の技術者一覧

No.	氏名	年齢	現住所	最終学歴	Python経験年数	職種	現在のプロジェクト	主なPython関連スキル	資格
1	T.N	37歳	東京都渋谷区	大学院 理工学研究科	~10年 (2016年4月~)	MLエンジニアリード	大規模推薦システム (日間1億円推論)	PyTorch, TensorFlow, Scikit-learn, BERT, AWS SageMaker, MLflow, Kubernetes	AWS ML Specialty, Google ML Engineer
2	A.Y	34歳	京都府京都市	大学院 情報科学研究科	~8年 (2018年4月~)	データエンジニアリード	データレイクハウス設計・運用 (日間5TB)	Apache Spark, Airflow, Kafka, Flink, dbt, Databricks, AWS	AWS Data Analytics, Databricks Certified
3	M.O	33歳	神奈川県川崎市	国立大学大学院 基礎理工学研究科	~7年 (2019年4月~)	機械学習エンジニア	AI画像認識モデルの実用化・MLOps基盤構築	PyTorch, TensorFlow, Scikit-learn, OpenCV, AWS SageMaker, MLflow, Docker	AWS ML Specialty
4	A.I	29歳	京都府京都市	国立大学大学院 情報学研究科	~4年 (2022年4月~)	データサイエンティスト	広告効果予測モデルの構築・データ分析	Pandas, NumPy, Scikit-learn, LightGBM, BigQuery, Tableau	統計検定2級, G検定
5	Y.U	26歳	東京都板橋区	私立大学 工学部 情報科学科	~3年 (2023年4月~)	バックエンドエンジニア	予約システムのバックエンドAPI開発	Django, Django REST Framework, PostgreSQL, Docker, unittest	Python 3 エンジニア認定基礎試験

補足

- T.N, A.Y, M.O はPythonを主力言語として10年~7年の経験があり、機械学習/データエンジニアリング領域での実績が非常に豊富です。
- A.I はデータ分析・機械学習領域でPythonを4年間使用。
- Y.U はWebバックエンド (Django) 開発に特化しており、Web系Python案件に最適です。

※ Goバックエンドエンジニアや若手データアナリストもPythonを使用していますが、前者はPythonがサブ言語、後者は「Python(基礎)」レベルのため除外しています。

ctrl esc to focus or unfocus Claude

+ [] | README.md

</> Edit automatically

