

ChatBrid:

■専ら社内知識とローカル情報だけで対話します
■1000本超のマニュアル群など大規模ナレッジ対応
■ハルシネーション率１％未満！（本稼働実績）

https://metadata.co.jp/services/chatbrid.html

ChatGPTを取り巻く課題

主な課題：
「情報が正確ではない可能性がある」
「専門性の高い質問には答えられない」
「フィッシング等への悪用可能性」
「悪性コードの作成に使われる可能性」
「著作権対応がグレー」
「入力内容が学習され第三者に漏洩する可能性」

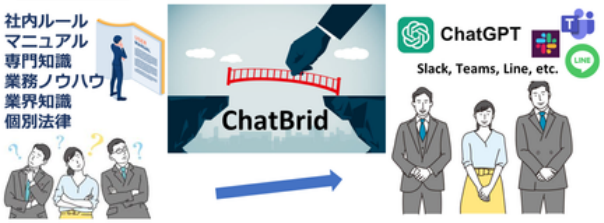




ChatBridが解決します

ChatBridとは

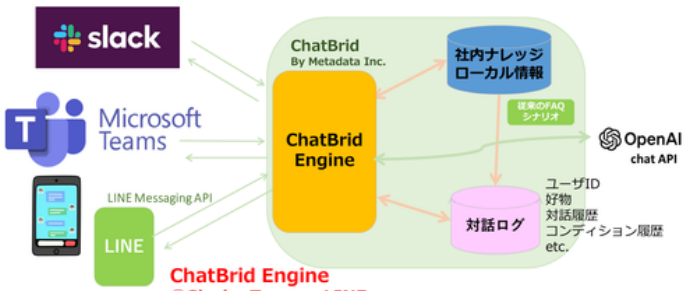
- Knowledge ManagementのDX化
- 社内、組織内に溢れている、固有のローカルな知識、情報を編集し、独自の自然言語処理で選考。ChatGPTに結びつけることによって、正確、適切にわかりやすく回答する支援ツール



https://metadata.co.jp/services/chatbrid.html

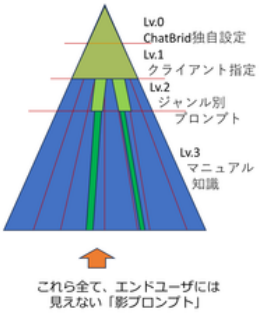
ChatBridの概要

～RAG命名より3か月先行。さらに高度化



ChatBrid Engine
①Slack、Teams、LINE...
使い慣れたフロントエンドでご利用可能
②Knowledge Management
独自の自然言語処理で高精度！ 人間回答者より分かりやすく好評！

何が ChatGPT API に送られるか？



これら全て、エンドユーザには
見えない「影プロンプト」

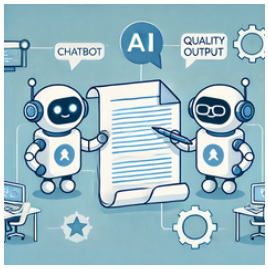
- 左記、緑部分（質問文との類似検索で上位ランク入りしたナレッジ）＋各部をつなぐ文構造
- 事例：
  - 20ジャンル、数100本のマニュアル類から約5000本のナレッジ素片をマークダウン形式でChatBridに取り込み
  - 質問文との関連度高い上位30本、最大5千～5万文字をヒットさせる
    - うち緑色部分がデータセットプロンプト
  - 4階層プロンプト＋つなぎ文による指示＆内容を分かり易い言葉で表現してChatGPTが回答

特許出願済

このデータセットプロンプトにより、速い、安い、巧い（高精度）を実現！

ChatBrid導入のご提案プロセス

- 社内で活用されていない（共有されていない）情報、社内サポートの問題点等の現状把握
- 導入後の効果予測
- お見積作成（期間、費用）
- 実施
  - 全体構成
  - 情報、データのマイニング
  - テスト
  - 運用開始
- 保守、メンテナンス



実施フェーズと参考規模の費用例

- Phase 0 PoCフェーズ
  - 平均10数ページのWord, Excel, Powerpointマニュアル10本程度。
  - 改善&マークダウン化、ChatBrid構築、精度評価:3～4w
  - 費用：お問い合わせください
- Phase 1 本番：知識開発フェーズ
  - 規模、情シス状況に応じNotePM等周辺ツールを適宜併用
  - マークダウン化&マニュアル改善の虎の巻をカスタマイズ、提供
  - 想定質問のタイプを共同で考案。定量評価用の質問文セット作り
  - 費用：お問い合わせください
- Phase 2 導入初期重点サポート・改善フェーズ:
  - 定常運用フェーズのサービスに加え、評価フィードバックに応じてプロンプト等の改善、サポート
  - 費用：お問い合わせください
- Phase 3 定常運用フェーズ:
  - 費用：ChatBridライセンス25万円/月～+ ChatGPT API利用料

Metadata

メタデータ代表 野村直之「**RAG博士**」による**自然言語処理**と**ナレッジマネジメント**の40年のノウハウを生成AIに活かした最強のナレッジ活用サービスを提供します。

お問い合わせ：**https://metadata.co.jp/contact.html**

ハイエンドRAG製品ChatBrid
Retrieval Augmented Generation

ChatBrid活用事例1
社内マニュアル全て+α から答える総務ボット

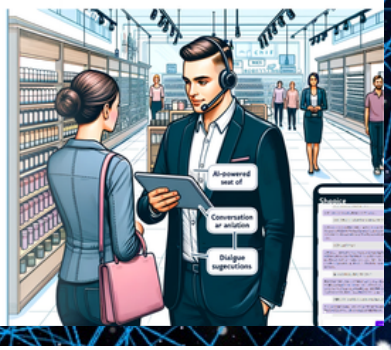


- 大手オンラインWebサービス提供社様
  - 各種ハルシネーションを１～２桁低減
    - 社内知識のみから回答
  - 大規模な知識体系を快適管理
    - 数10～数100の「データセット」に分割
    - 1000超の専門知識をデータセットごとのプロンプトに分割
      - GPT側の精度、応答速度を向上。課金も削減可能
  - 大規模化で精度急降下するRAGの欠点を解決
  - 標準搭載の評価ツールで1000件級の評価を毎日実施
    - 回答や質問との関連性をナレッジ上に多色付けて瞬時に把握
  - 単語だけでなく意味による関連・類似を事前把握してからChat APIに送付
    - 30数万単語にあらかじめ付与済の、1万種類5階層の意味分類コードにより、社内マニュアルにない言葉で質問されても高精度で回答

ChatBrid活用事例２
店舗での模範顧客対話の生成、採点!

・顧客プロフィール、既購買歴、嗜好、生活習慣、今回の来店目的を入力 (ゆくゆくは顔画像認識・音声認識も参考に!)

- **接客で外せない内容を網羅し、お買い上げに直結**する対話を生成
  - 高々数10種類の対話例を影プロンプトに搭載
  - 過去の購買履歴ごとに商品名、数量、利用目的、クレーム履歴等の知識を搭載
  - 影プロンプトに例示していない顧客プロフィールや商品についての、模範的な接客対話を生成
- AIによる生成結果を範として新人販売員の接客対話の内容を採点、評価



ビジュアル類似検索で楽々評価改良
数万件の中に質問内容が**無い**ことさえ数秒で判明！

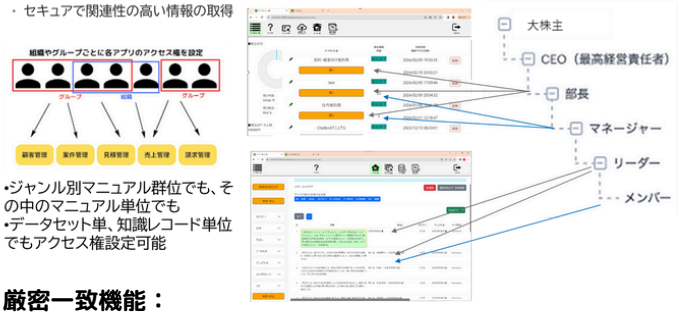
・**社内ナレッジ**、マニュアルには、抜け、矛盾、改版漏れ等が、通常山ほど残っているのを発見し改善へ



悪魔の証明ができるツール！

知識アクセス権制御

- 質問者の所属や資格を確認し、アクセス権限内でのナレッジのみ参照し回答生成へ。
  - 企業のセキュリティとコンプライアンスに重要
  - セキュアで関連性の高い情報の取得



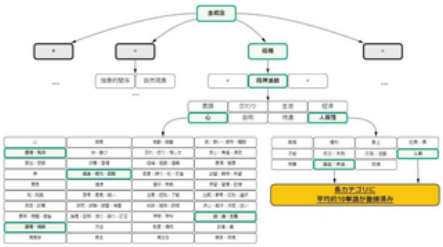
•ジャンル別マニュアル群位でも、その中のマニュアル単位でも
•データセット単、知識レコード単位でもアクセス権設定可能

厳密一致機能：

薬品名10mg vs 1000mg等はLLMは混同しがちだが、それは致命的！
→ 指定した属性の対象は文字列が厳密に一致させるRAG機能も提供して解決

質問文を階層シソーラスで展開
単語だけでなく意味の類似性で候補を拾う

～30数万単語を5階層約1万種の意味に分類済



- 一般の質問者はいろんな言葉、言い方で表現
  - ChatGPTはそれらの意味を解釈してくれるが、GPT側に送付前の普通のRAGは表現の違いに対応できない
- **ChatBrid**なら、程よく類義語に展開可能！
  - 影プロンプトに、本命回答の素材を含める確率UP！
- 対義語（反対語）も含まれる
  - →否定表現でも関連性をカバー

類義語ベクトル要素の重みはパーセント刻みでコントロール可能

マークダウン構造を自動解析し知識レコードに分割

例：NDAの雛型を条文ごとに自動分割してナレッジ化
「区切り文字」を # ## ### #### とすることで、これらの区切りで知識レコード（チャンク）を生成。



上位見出しを複製・追加するコンテクストフリー化オプション搭載

コンテクストフリーONにすると、「見出しと本文」の組からなる１知識レコードは次のように構成される：

見出し：第１条（定義）…　≪ 第２章 生成AIと専門知識 ≪ 第一部 生成AI 応用概説 （≪ ○○2024年度報告書）

これら上位見出しがマークダウン構造から自動補完される

質問の「言い換え」「翻訳」機能による再現率向上

独自ノウハウにより精密にコントロールされた言い換えによる展開例：
元の質問「当社はAPIサーバの仕様を変更しないとありますが、利用者との合意があった場合はどうなりますか？」

- 当社はAPIサーバの仕様を改訂しないと述べていますが、利用者との合意が存在する場合はどうなりますか？
- 当社はAPIサーバの機能を変更しないと記されていますが、ユーザーとの合意があった場合はどう扱われますか？
- 当社はAPIサーバの仕様を変更しないとされていますが、利用者との合意があれば、どういう扱いになりますか？
- 当社はAPIサーバの仕様を変更しないとのことですが、利用者との合意があった時はどうなるのでしょうか？

翻訳による展開例：
元の質問「当社はAPIサーバの仕様を変更しないと定めていますが、利用者との合意が必要な場合はどのような手段で行われるのでしょうか？」

- We have established that the specifications of our API server will not be changed, but what methods will be used if an agreement with the users is necessary?
- 我们已经规定不会更改我们的API服务器的规格，但如果需要与用戶达成协议，将采用什么方法？
- Nous avons établi que les spécifications de notre serveur API ne seront pas modifiées, mais quels moyens seront utilisés si un accord avec les utilisateurs est nécessaire ?

他方式と違い、速度低下なし！
ランキングを繰り返すRAG Fusionやリランキングと違って最初の1回のLLM呼び出しで完結 → 精度（適合率）は同等以上

博士論文が次世代高精度RAGの仕様書に！
… 認知科学的辞書と情報検索・要約に関する研究 (野村2002)



目次

第１章 序論
第２章 用言格フレーム知識の記述、分類のための アーキテクチャとその検証
第３章 大規模電子化辞書とイディオム的語彙、表現の知識表現
第４章 脳内辞書の制約能力と文法との接点 ～格フレームの交代現象を説明し中間言語の理論的基盤を強化する普遍辞書理論
第５章 情報検索エンジンの研究開発による普遍辞書理論の有効性の検証 →RAG改良の仕様書！
第６章 文の分割による長文短縮機能を備えた文章要約システム
第７章 知識管理の現場における検索・要約技術の活用 ・評価と今後の課題
謝辞
付録1 …

第５章 情報検索エンジンの研究開発による普遍辞書理論の有効性の検証
5.1 言語学的知見、自然言語処理の高度化と相性の良いヴェクトル空間法
5.1.1 ヴェクトル空間法とそのターム・スコアを与えるtf*idf法
5.1.2 複合名詞句間照応(浅い文脈解析)によるデータ疎問題の解決と精度向上 ～タームの定義に言語学的知見、形態素解析結果を取り込む
5.2 複合名詞句間照応ルールに普遍辞書理論の知見を導入
5.2.1 非語彙的なNvV複合語に限った、名詞句と単文との照合
5.2.2 普遍辞書理論の知見による副作用の少ない精度向上効果
5.3 類似検索と関連性フィードバック
5.3.1 類似検索の特徴と長所
5.3.2 ConceptBase Searchの特徴のまとめ
5.3.3 なぜ関連性フィードバックが必要か？
5.3.3.1 事前に目的物を特定できない／表現できない問題
5.3.3.2 探索空間に含まれる予期せぬ不正解文書と関連性フィードバック
5.3.4 類似検索における新しい関連性フィードバック・パラダイム
5.4 ピンポイント検索：再現率重視から適合率重視の情報検索へ
5.4.1 主要な格要素の係受け解析による高精度文書検索
5.4.2 高速性、使用リソース、のバランスを維持しつつ係受け検索を実現
5.4.3 係受け検索の評価と考察