高精度大規模RAGエンジン ChatBrid の技術的優位性

- LLMは大規模化で精度向上するがRAGは通常精度低下
- 様々な技術的、内容的工夫で遂にナレッジマネジメントの夢が実現！

独自のベクトルストレージとTF*IDFの重み調整

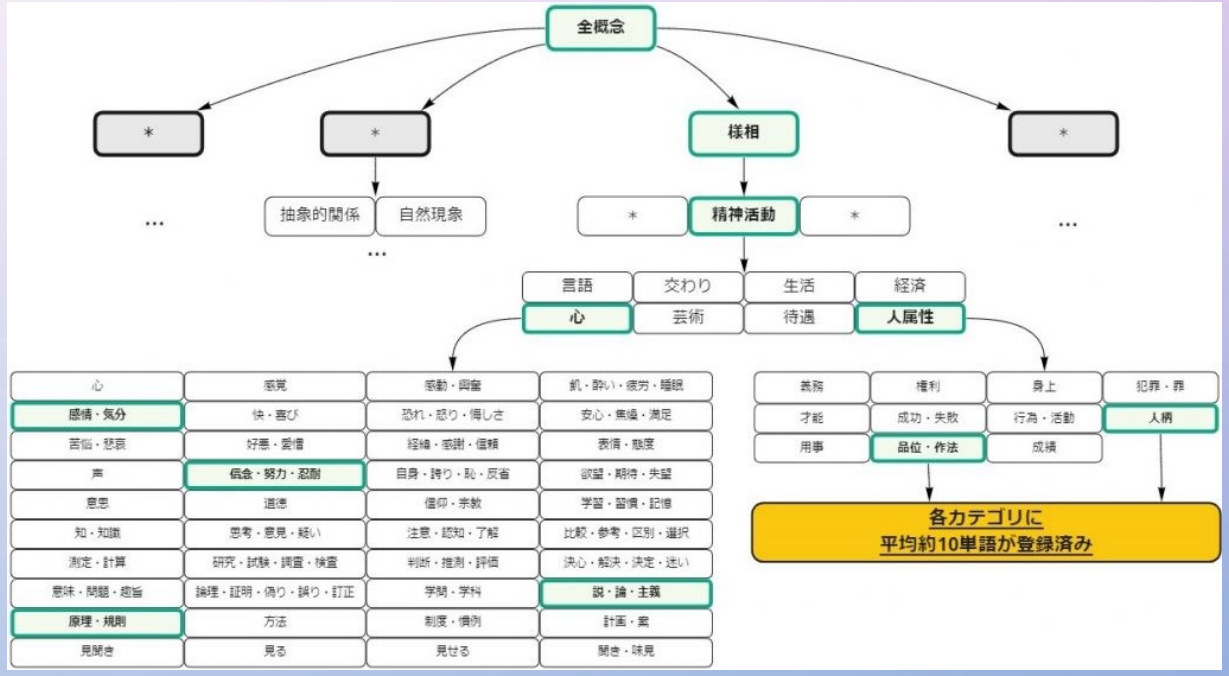
- 独自のベクトルストレージにより、TF*IDFの重みを自在に調整
- 日本語のSTOP WORDS（不要語：の、こと、もの、する 等の重みをゼロに）を数百語登録
- 単語分割を精緻化。複合語全体と構成語を共にベクトルの要素とし検索精度を向上
  - **沖縄基地問題 vs 沖縄米軍問題 が 67%類似**
  - **3.3mg vs 3月3日 が類似しない**

(text:型)^3.0 text:溶解 (text:倉庫)^2.0 text:取引 text:t (text:除く)^2.0 (text:手形)^13.0 (text:日)^15.0 text:ラード text:鉛 (text:買)^3.0 text:ビートグフ text:段ボール (text:名古屋)^2.0 text:雑貨 (text:鋼板)^2.0 text:油脂 text:営業 text:テレフタレート (text:製品)^2.0 (text:米)^2.0 (text:卸)^3.0 text:叩 text:天 (text:ポリエチレン)^2.0 (text:繊維)^2.0 ...

経済特集：日本市場における多様な素材と製品の最新取引動向
【名古屋発】近年、名古屋を中心とした日本の工業市場では、多岐にわたる素材と製品の取引が活発化しています。特にポリエチレンテレフタレートやポリエチレンなどのプラスチック素材、鉛、油脂といった化学製品の需要が増加しており、これに伴い、営業戦略も急速に変化しています。
最新の市場データによると、名古屋の倉庫ではこれらの素材が大量に取り扱われ、特にポリエチレン製の段ボールやビートグフと呼ばれる特殊な繊維製品の在庫管理が重要視されています。これらの製品は、国内外の雑貨市場での需要が高く、日々の取引量も増加傾向にあります。...

5階層シソーラスの活用

- 30数万語の5階層シソーラスを活用し、単語の意味カテゴリをベクトルの要素に追加
- 意味カテゴリをベクトルに追加
- 質問文中に、マニュアルに無い類義語のみが使われている場合でも正しい知識がヒット
  - 一般の質問者は様々な言葉、 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かった
  - **CHATBRID**なら、程よく類義語に展開可能！
  - 対義語（反対語）も含まれる



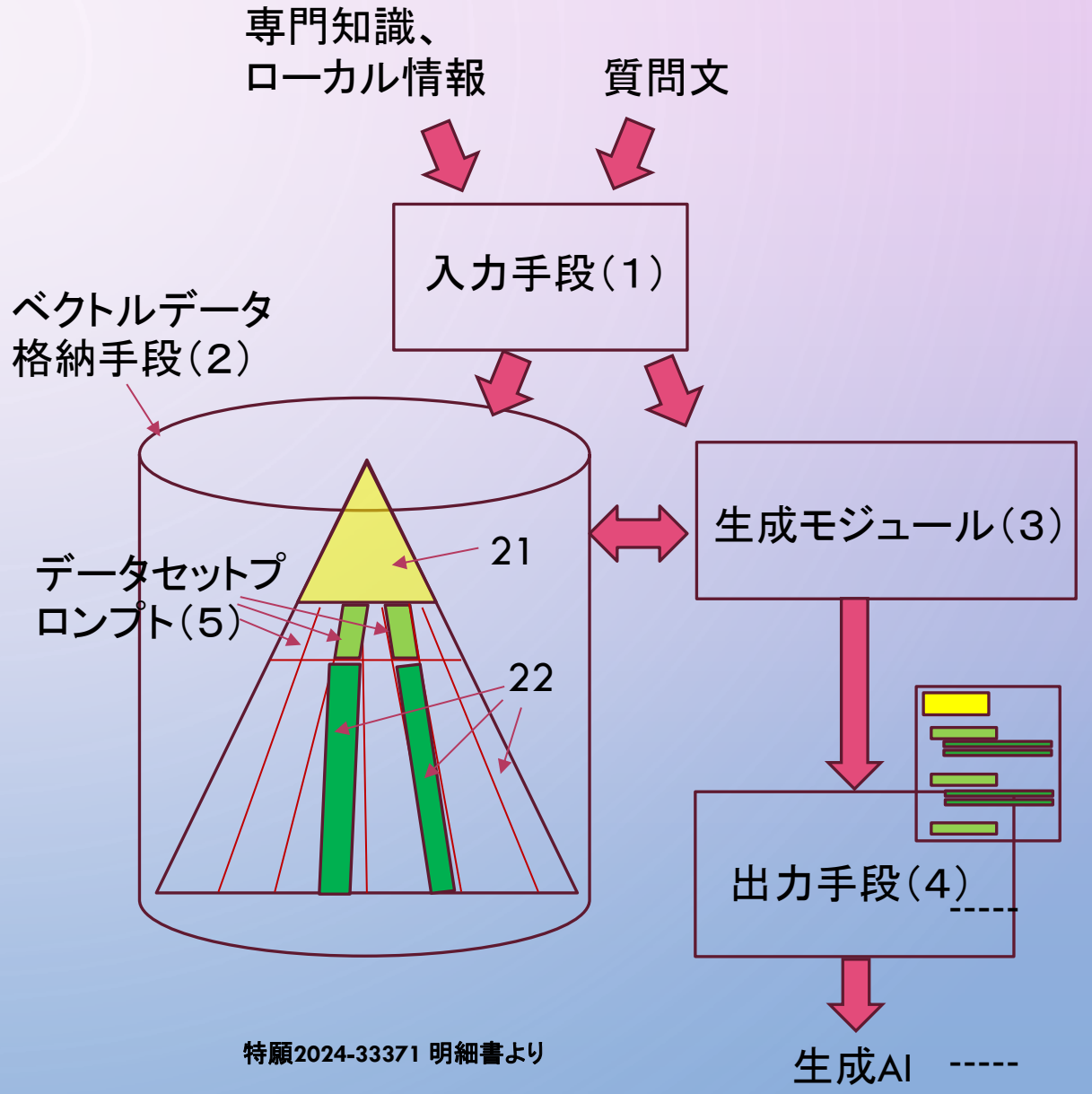
ベクトル検索結果のビジュアル化
～知識デバッグのための視覚フィードバック

- 自前のベクトルストレージにより、回答不能や不適切な回答のベクトル検索結果を色付け表示
- **ビジュアル類似検索でAIの回答を秒速で評価**
  → ナレッジ自体の有無、品質を評価、改良
- マニュアルの記述が不適切だった場合の即座のビジュアル確認と修正が容易。

- マッチした単語と意味コードを強調表示
- 不適切な回答に対する即時の視覚フィードバック
- MARKDOWNをその場で修正 → 即検索に反映

特許技術による影プロンプト送出量の削減&高精度化

- ナレッジをジャンル分けし各データセットに分割登録することにより影プロンプトの平均量を削減。
  - RAG内の検索でヒットしたデータセットに付随する影プロンプトのみをLLMに送信。
  - これによりLLMのレスポンスが「速い」「安い」「旨い（高精度）」
    - 高速な応答と低コストのLLM利用
    - 大規模化しても高精度を維持
    - 通常1専門分野に特化するRAGをマルチ専門家に
      - 質問文に関連の深い上位ナレッジに付随のデータセットプロンプトのみをLLMに送出
      - コンシエルジェ、総務的な機能が可能に
    - 数百本、数千本のマニュアル、数万ページ、数10万件のナレッジを取り入れられる。



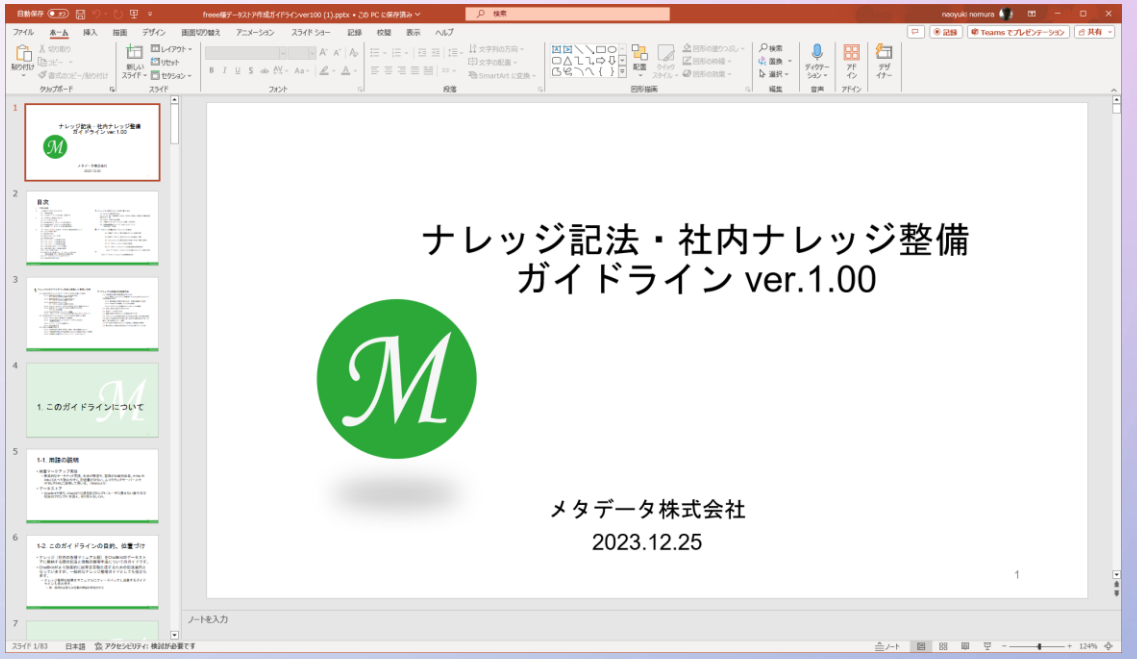
特願2024-33371 明細書より

データセットや各種絞込み条件下の単語分布を目視

- 高頻度の未知語から、用語説明を付すべき社内専門用語などを素早く発見

知識管理のための包括的なマニュアル100数10ページ

- DOCX, XLSX, PPTXをLLMが理解しやすいマークダウンに変換し、知識管理を効率化。
- 100ページ超の効率的なMARKDOWN変換マニュアル
- DOCX、XLSX、PPTXフォーマットに対応

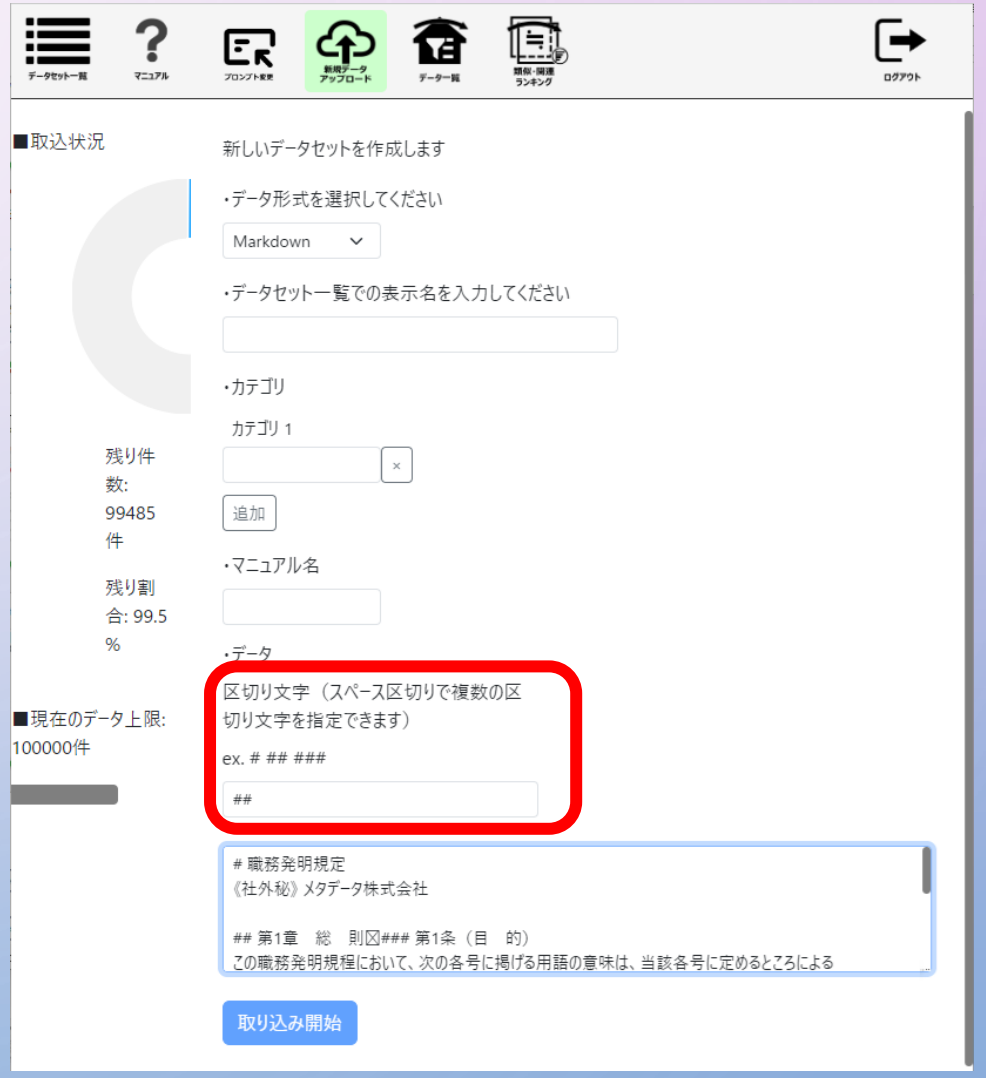


目次

自動レコード分割
～マークダウンの見出し階層指定に基づきチャンキング

職務発明規定
《社外秘》メタデータ株式会社

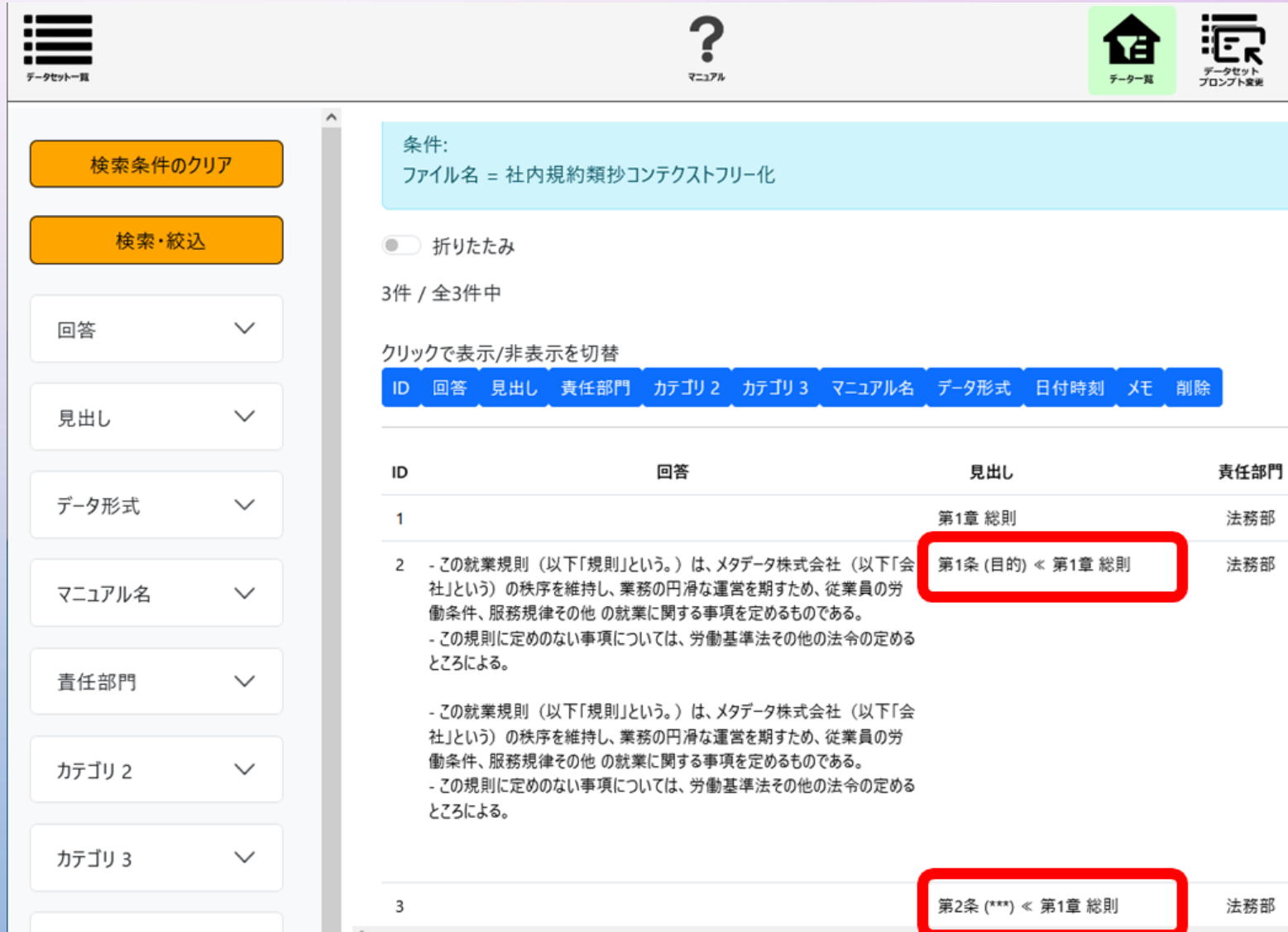
第1章　総　則
第1条（目　的）
- 　この職務発明規程において、次の各号に掲げる用語の意味は、当該各号に定めるところによる
第2条（用語の定義）
- 　この職務発明規程において、次の各号に掲げる用語の意味は、当該各号に定めるところによる。
  -- 職務発明　その性質上会社の業務範囲に属し、かつ、その発明をするに至った行為が会社における従業者等の現在または過去の職務に属する発明として第5条に基づいて会社が認定したもの。
  -- 発明者　発明をした従業者等。
  -- 従業者等　期間の定めの有無を問わず会社が雇用する者および会社の役員。



知識粒度を小さめにして高精度化しつつコンテクスト情報(上位見出し)を自動付加しLLMの理解を促進

- 知識粒度(レコードサイズ=文字量)がバラつき過ぎると精度低下
- 特に巨大レコードには複数トピックが存在し質問文に無関係なのに関係ありげな記述が混入し精度低下
- 個々のレコードは1パラグラフ程度が好ましいがコンテクスト不明になってLLMが理解失敗することも

- 小粒度でのコンテクスト維持のために上位見出しを自動付加！



厳密一致遵守機能と属性による知識絞り込み

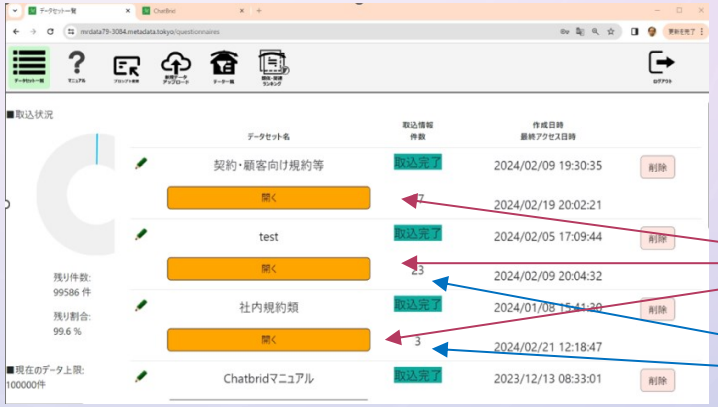
- 指定した文字列との厳密一致遵守(EXACT MATCH)を条件にすることで材料名の取り違えなどの致命的な誤りを防止。
- 属性、法律・条令の有効期間、適用地域等による知識の厳密な絞り込み。

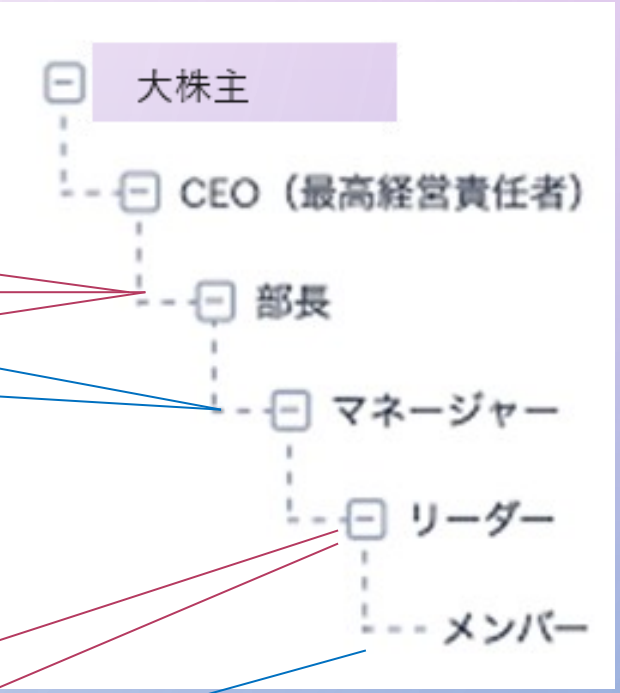
材料ID	材料名	純度	供給元	在庫量 (kg)	備考
Cu-101	電気銅 (C1100)	99.99%	A社	1000	高純度銅
Cu-102	りん脱酸銅 (C1220)	99.90%	B社	500	少量のりんを含む
Cu-103	無酸素銅 (C1020)	99.96%	C社	750	酸素含有量が極めて低い
Cu-104	タフピッチ銅 (C1100)	99.90%	D社	1200	一般的な純銅
Zn-201	高純度亜鉛 (Z1100)	99.99%	E社	600	高純度亜鉛
Sn-301	高純度スズ (Sn99.9)	99.90%	F社	400	高純度スズ
Ni-401	電解ニッケル (N02200)	99.90%	G社	300	高純度ニッケル
Al-501	工業用純アルミ (A1050)	99.50%	H社	800	一般的な工業用純アルミ

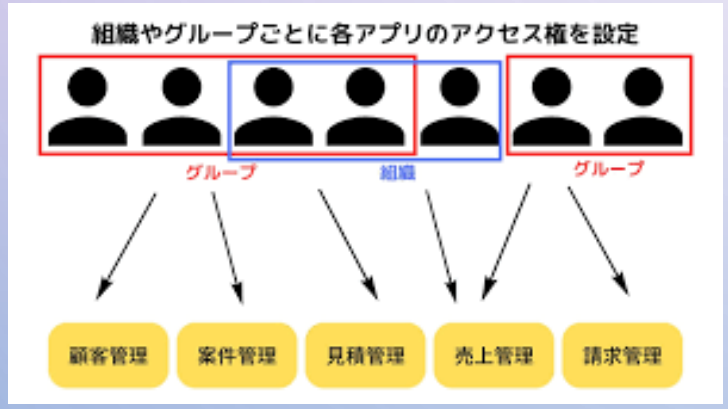
Cu-101とCu-102やZn-101が類似検索されてはならない。厳密一致遵守（Exact match）で絞り込み、該当レコードのデータのみを回答に反映させるべし。

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。
  - セキュアで関連性の高い情報の取得
  - 企業のセキュリティとコンプライアンスに重要







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

APIモードでの統合運用
～企業システムとの相互接続性

- CHATBRID全体がAPIモードで動作
- 各種企業情報システムと一体化して運用可能。
- アプリ構築環境MIIBOやDIFYと統合済。
  - シナリオベース、ワークフローベース

webhookによる連携（ChatBridのAPI）

miibo

Dify

…

API呼び出し

RAG DBの検索結果
＋
データセットプロンプト