

知識DBの自己増殖とは その1

2026年9月10日

株式会社礎デザインオートメーション

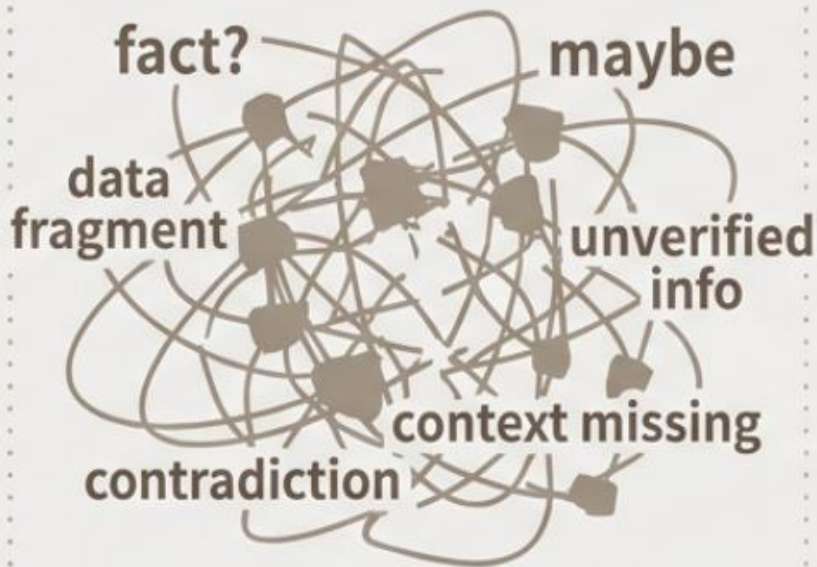
本書はあくまでも知識の自己増殖の考え方をまとめたものです。実際の運用には、誤りの自己増殖などいくつかの問題に対処する必要があります。

それら問題への対策や他の分野への応用などは順次紹介したいと思います。

知識の自己増殖と次世代知能インフラ： LLM×ナレッジグラフが導く自律型AI

AIの「もっともらしい嘘（ハルシネーション）」を克服し、
自ら正しく学習・推論する次世代のナレッジインフラを構築する。

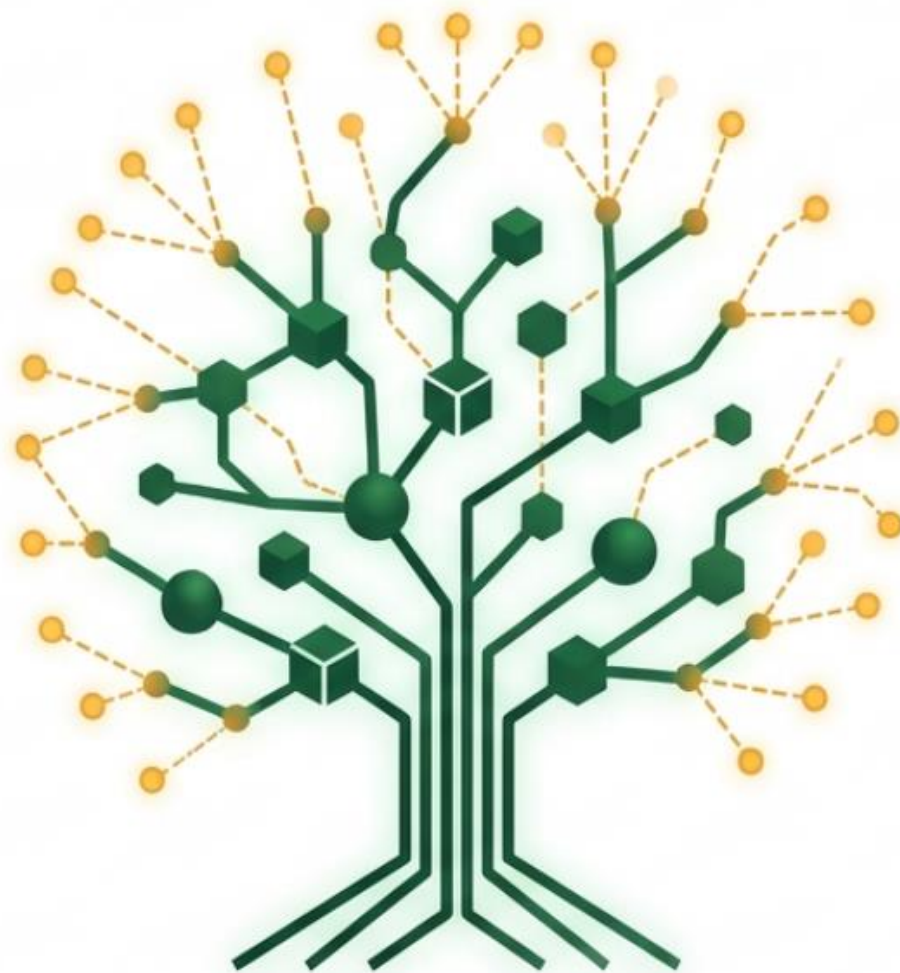
LLM単体



AI活用における最大の壁：
ハルシネーションとブラックボックス化

単なる「データの蓄積」から
「知識の自律的な拡張」へのパラダイムシフト

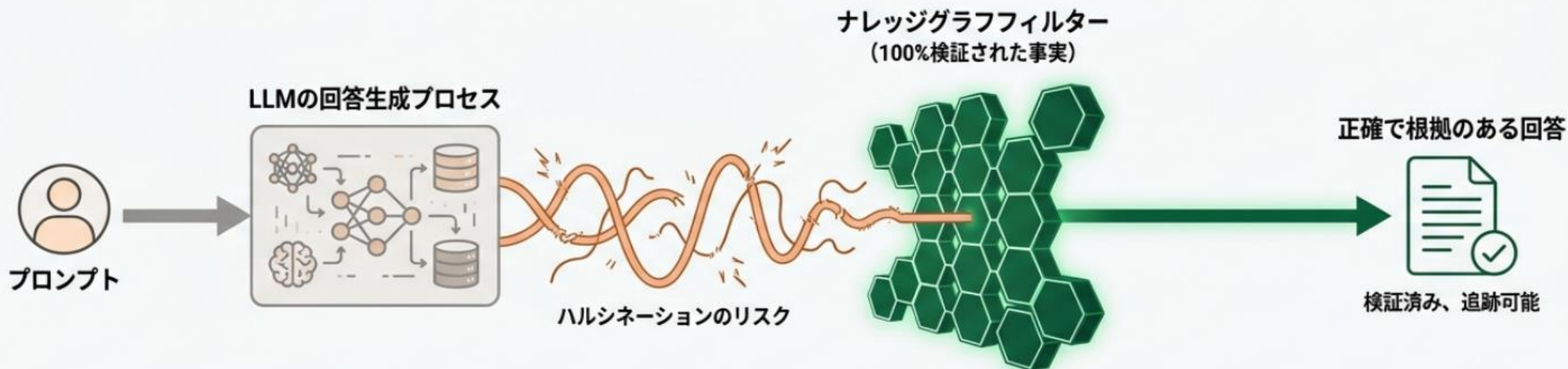
LLM + ナレッジグラフ



グラフDB、オントロジー、LLMの融合が
もたらすエンタープライズAIの最適解

LLMの限界とナレッジグラフの必要性

LLM単体では論理的な整合性を担保できないため、「100%正しい事実とルール」を管理するナレッジグラフが不可欠である。



LLMの弱点：確率に基づく文章生成によるハルシネーション（もっともらしい嘘）

LLMは大量のデータから確率的に最も尤もらしい単語を繋ぎ合わせるため、事実と異なるもっともらしい嘘（ハルシネーション）を生成するリスクが常に存在する。

説明責任の欠如：AIがなぜその結論に至ったか、根拠の遡求が困難

LLMの内部プロセスはブラックボックス化しており、なぜ特定の回答に至ったのかという論理的な根拠や情報源を明確に追跡・説明することが難しい。

ナレッジグラフの台頭：事実と関係性をネットワーク状に構造化し、確実な根拠をAIに提供

ナレッジグラフはエンティティ（物事）とリレーション（関係）をグラフ構造で管理し、検証済みの事実を明確に定義・接続することで、AIに確実な知識基盤を提供する。

RAGからの進化：単なる文書検索を超えた、文脈と論理に基づく高度な情報還元

従来のRAG（検索拡張生成）が関連文書を検索するのに対し、ナレッジグラフは文脈や論理構造を理解した上で、より高度で正確な情報をAIに還元し、回答の質を向上させる。

「知識の自己増殖」とは何か？

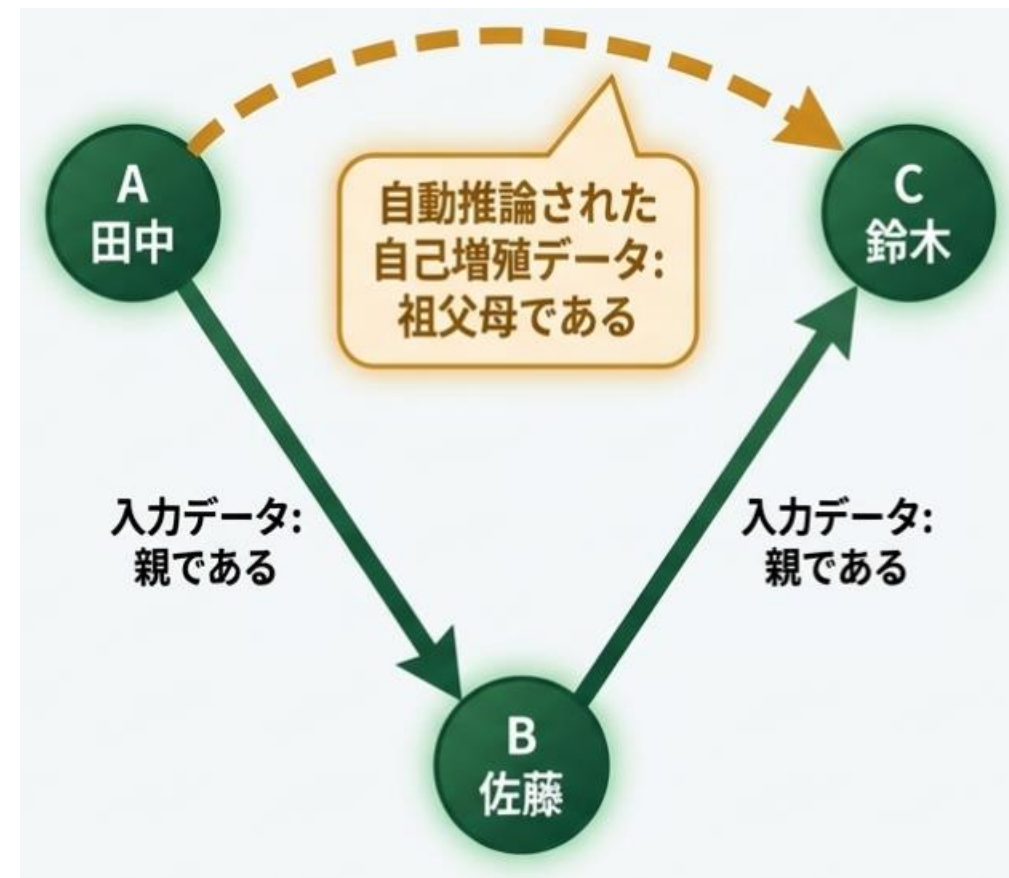
システムが既存のデータとルールを掛け合わせ、誰も教えていない「**未知の事実**」を自律的に生み出し続けるプロセスである。

情報の自律的獲得：AIが非構造化データから新たな事実を抽出し続ける

論理的推論による拡張：既存の事実とルール（常識）から、自動的に新しい関係性を導出

安全な結合：外部の知識ベース（オープンデータ等）とも矛盾なく融合

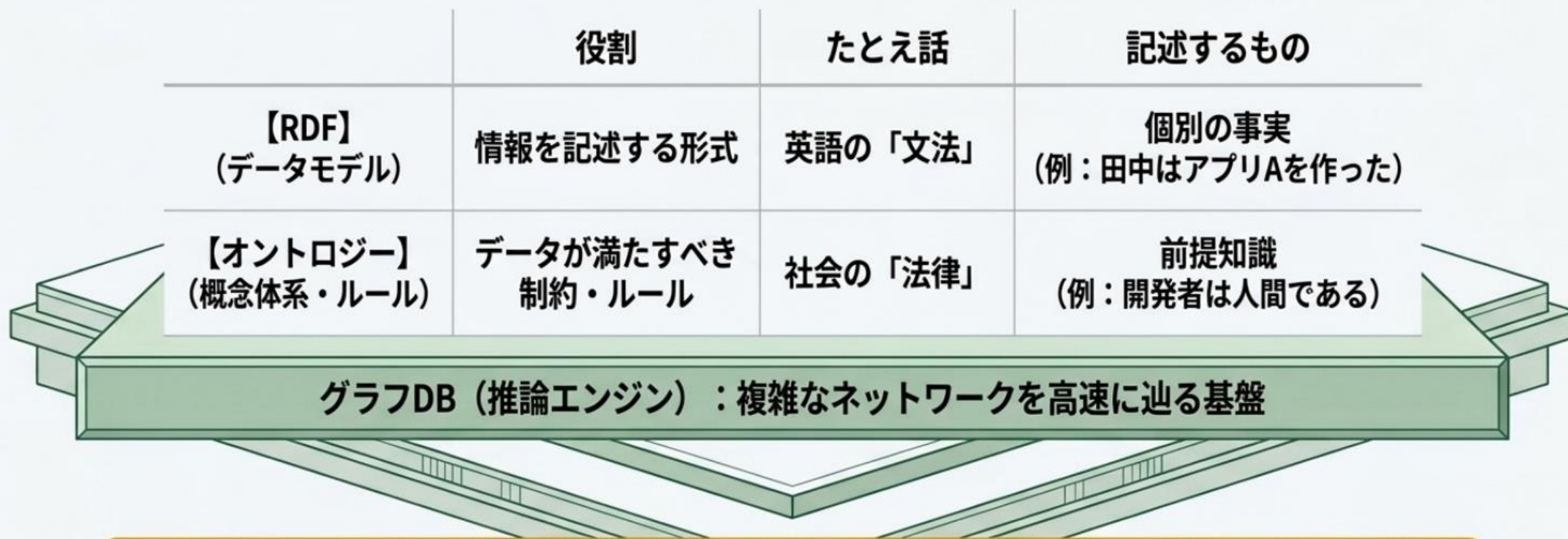
結果：データの入力数を超えるスピードで、知識ネットワークが指数関数的に拡大



RDF (トリプル表現)で、すべての情報(データとルール)を「**主語- 述語- 目的語**」の最小単位で記述。

「データ」と「ルール」の明確な分離

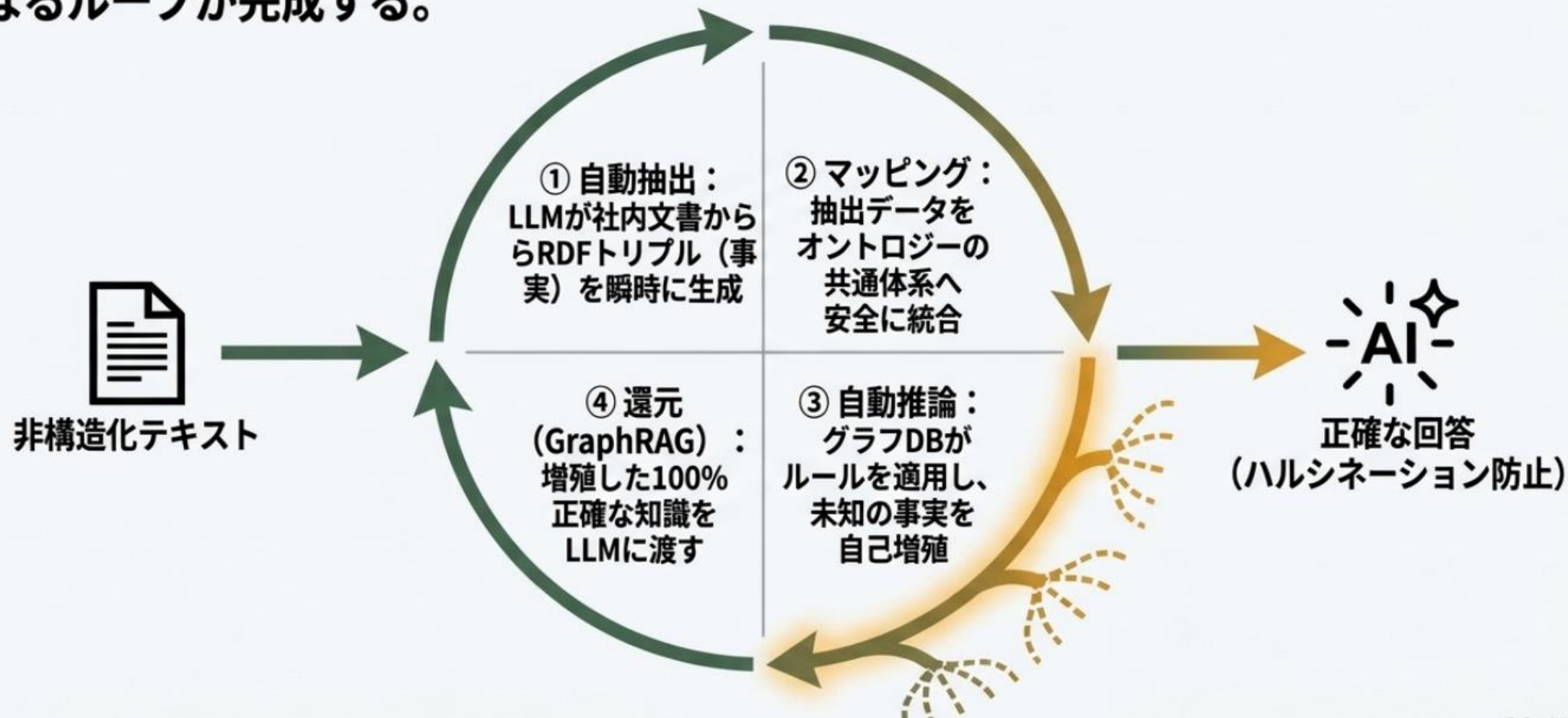
単なるデータを「意味のある知能」に変えるため、文法（RDF）と法律（オントロジー）を切り離して管理する。



分離のメリット：矛盾データ（バグ）の排除と、システム拡張時の圧倒的な柔軟性

4ステップで回る「自己増殖サイクル」

LLMが事実を抽出し、グラフDBがルールに沿って推論・拡張することで、無限に賢くなるループが完成する。



製造サプライチェーンのユースケース

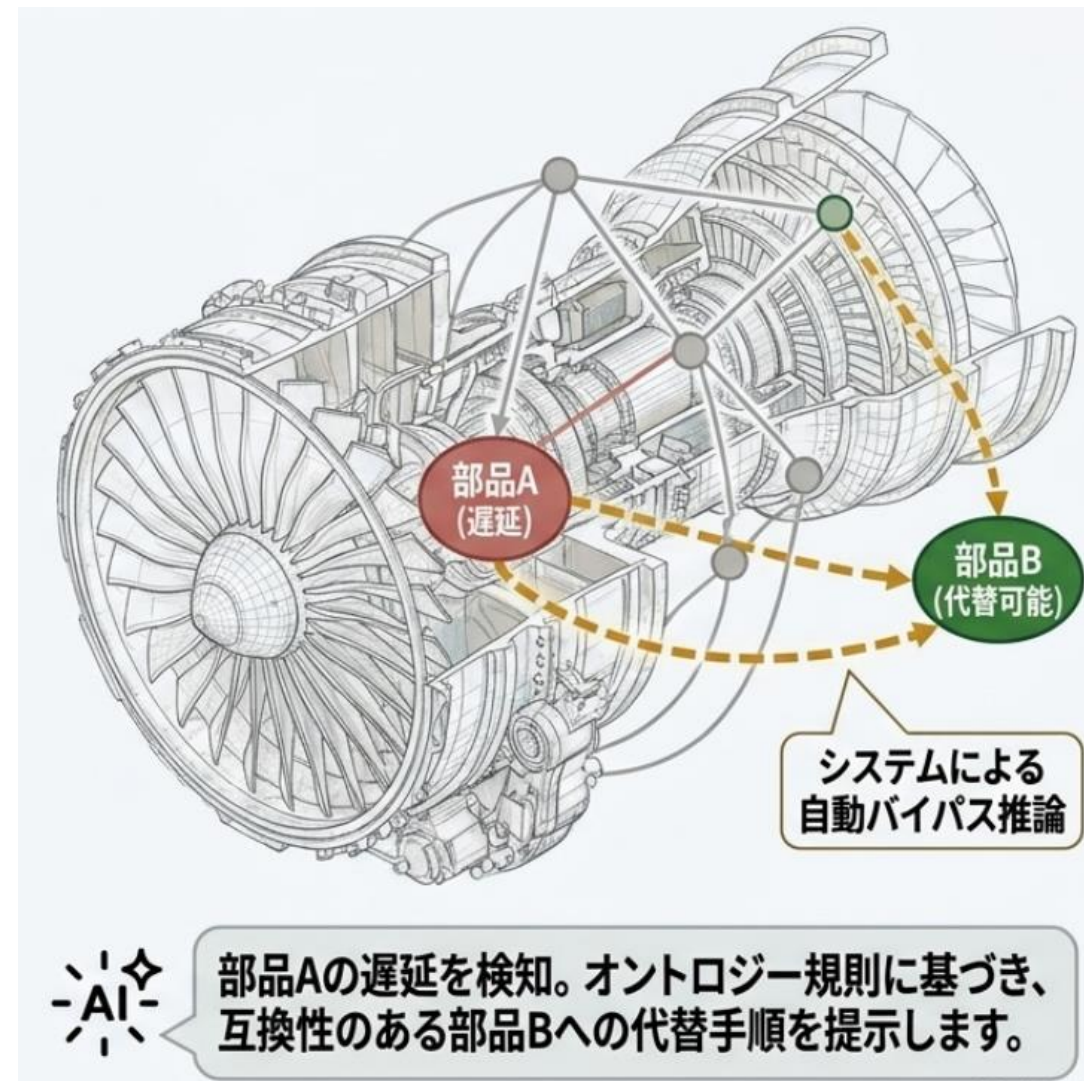
複雑な部品の依存関係と論理ルールをシステムが理解し、遅延影響の予測と代替案を自律的に提示する。

データの結合：CADデータ、BOM（部品表）、工場ログをRDF化して統合

知識の自己増殖：ルール（部品互換性）＋事実（納期遅延）＝推論（代替指示を自動生成）

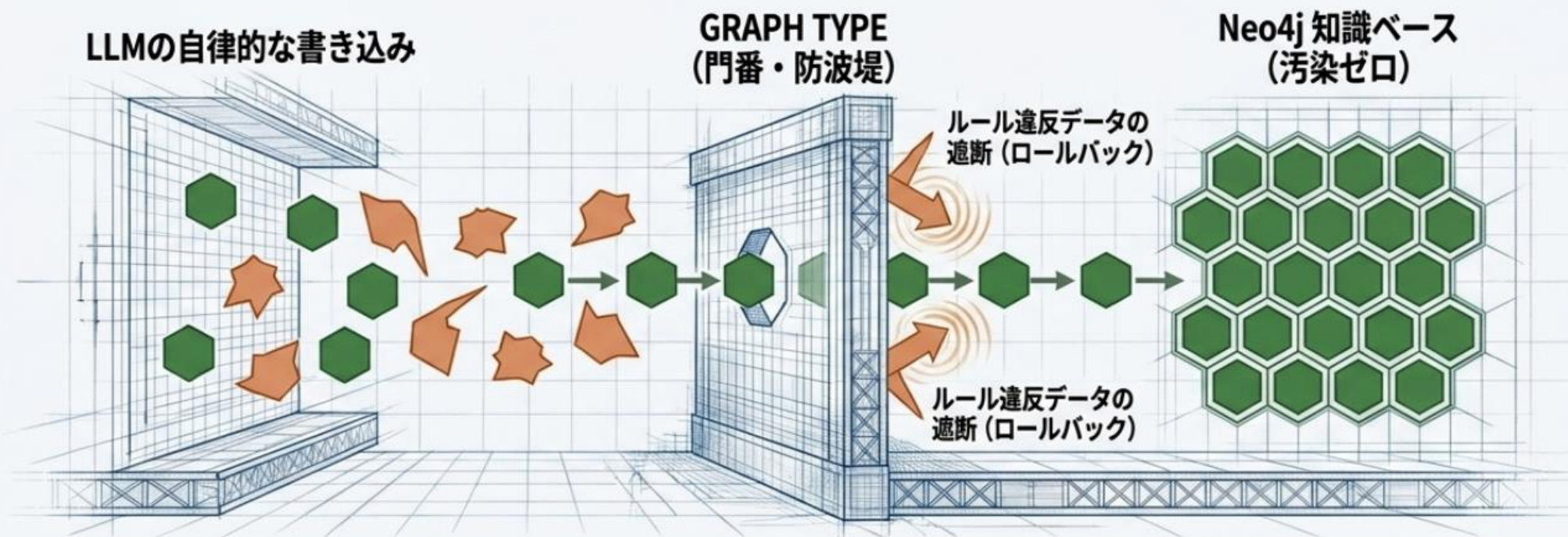
LLM連携：現場が「部品Aがない」と口頭指示するだけで、AIが最適な安全手順を即答

価値：属人的なノウハウの排除と、トラブル時のダウンタイム極小化



Neo4j「GRAPH TYPE」がもたらす安全性

Neo4jの最新機能「GRAPH TYPE」が、AIによるゴミデータの混入を防ぐ強力な防波堤と仕様書になる。



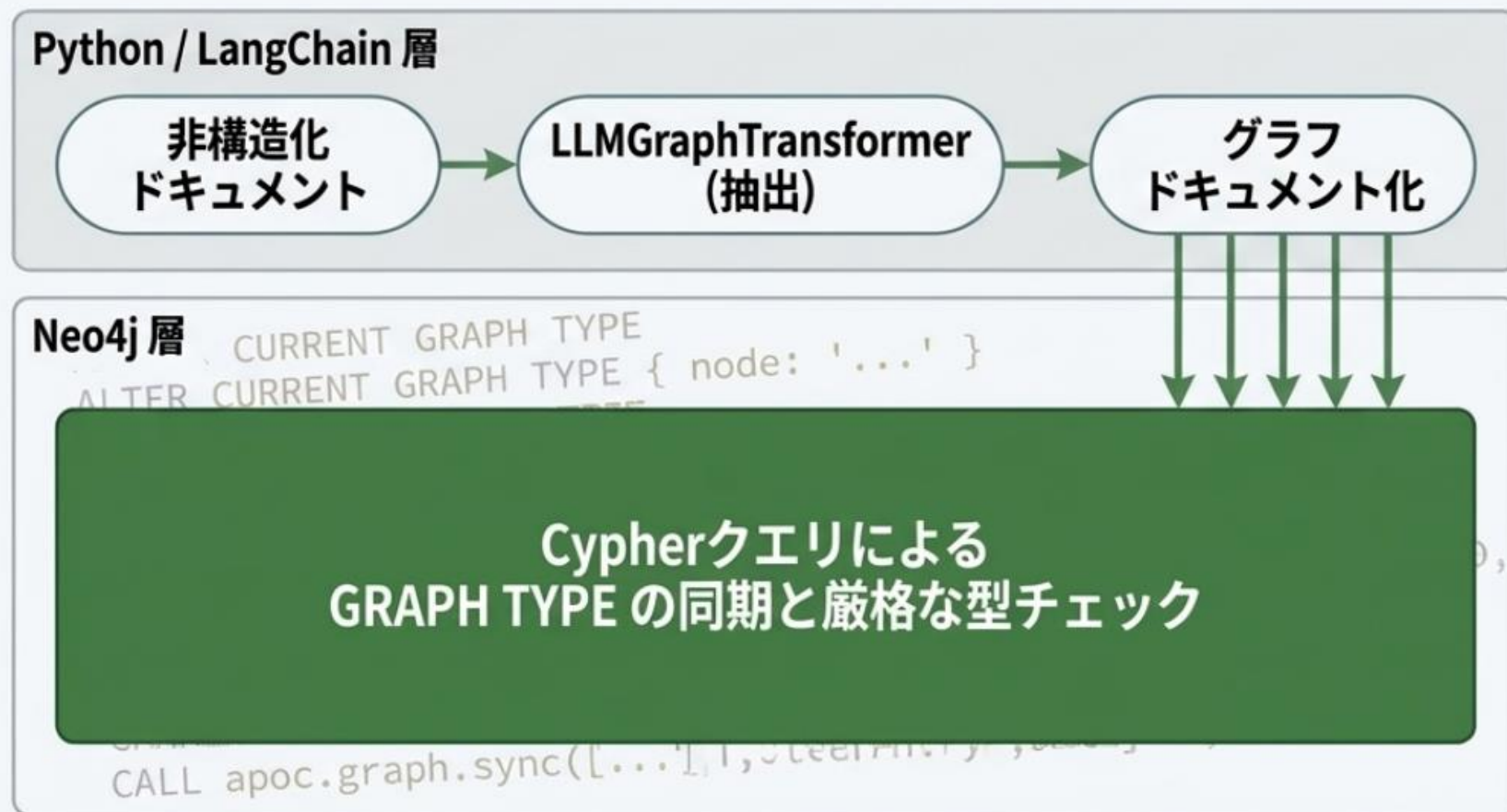
課題：AIの自律的な書き込みによる
データ汚染リスク

防波堤機能：データベース層での
「型」の強制。違反データを即座に遮断

柔軟なガバナンス：コアなルールは
死守しつつ、未知の概念の追加は許容

LangChainとNeo4jによる自律型パイプライン

LLMの柔軟な抽出能力とNeo4jの厳格な型制約を組み合わせることで、堅牢な自己増殖システムを実装できる。



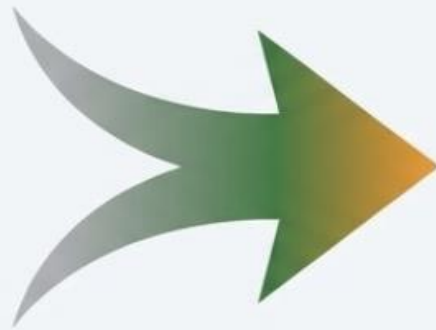
- SPARQL / Cypher連携：自然言語をデータベース言語へ瞬時に翻訳
- ゼロダウンタイム更新：システム稼働中のオンライン・オントロジー同期
- クレンジングの自動化：型のマイグレーションと不整合データの事前検知

次世代の知能インフラとして展望

単にデータを蓄積する「箱」から、人間のように自ら学び、考え、進化し続ける「知能」へ。



旧来のDB（静的な格納庫）



次世代の知能インフラ（成長し続けるナレッジグラフ）

トリプルストアによる、
極めてロジカルで高速な
エンタープライズ基盤

AI（LLM）の柔軟性と、
ナレッジグラフの厳密性の
ベストミックス

属人化していた「社内の
常識や専門知識」を、
拡張可能な企業資産へ転換

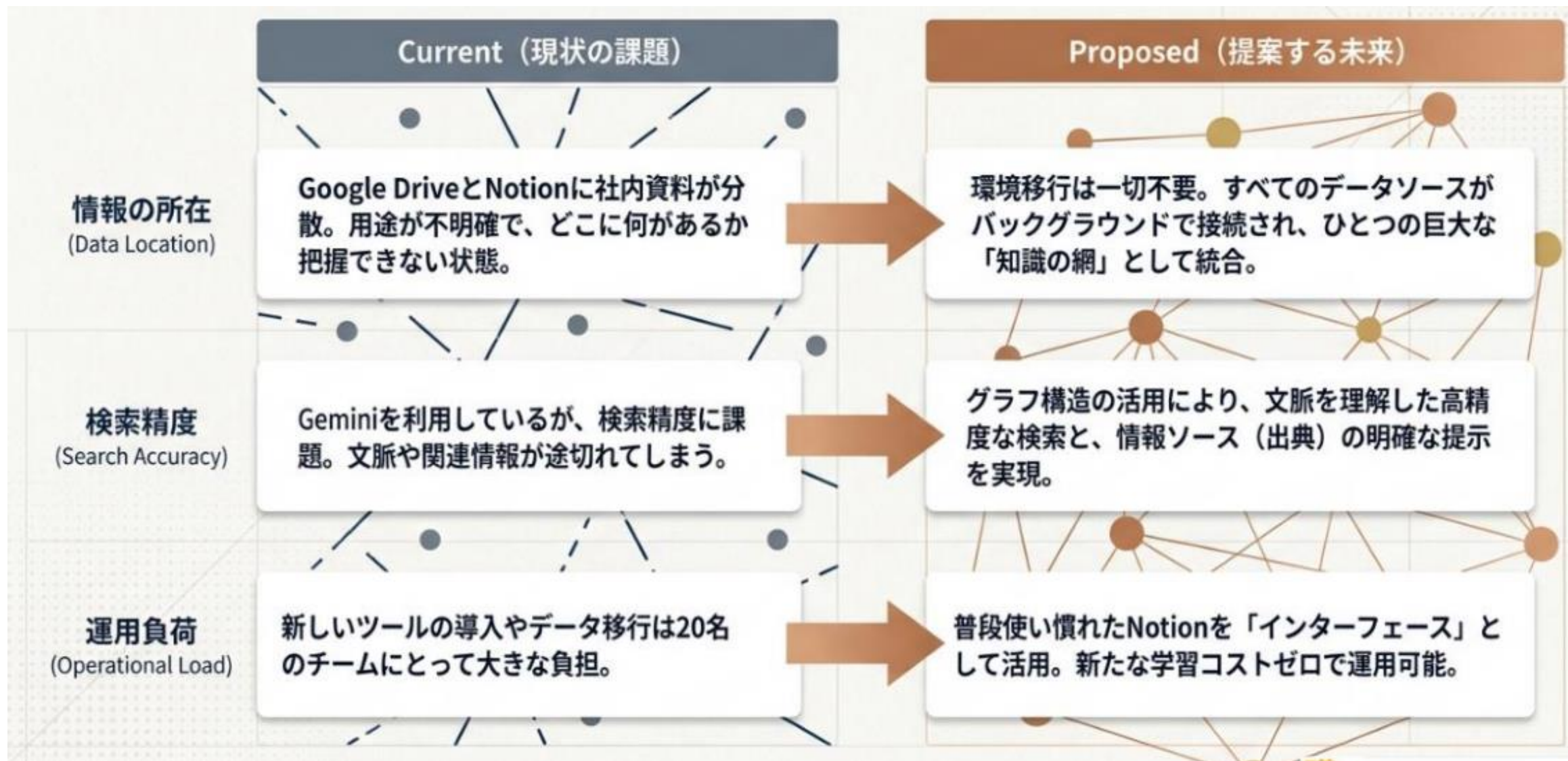
医療、金融、製造など、
高い信頼性が求められる領
域からビジネス変革を牽引

知識DBの構築の例

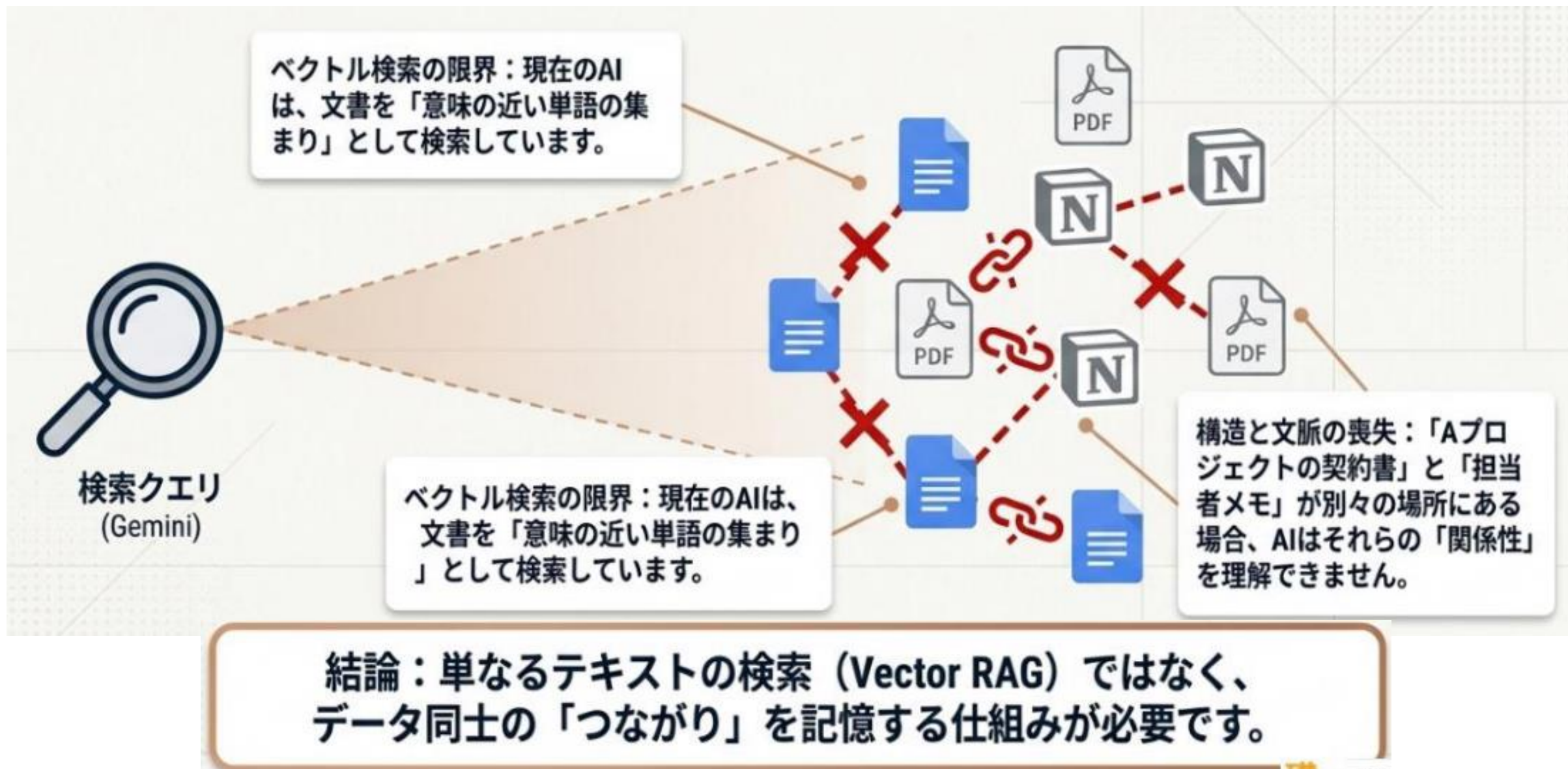
散在する情報から、ひとつの「生きた頭脳へ」

Notion X Neo4j X Gemini連携による次世代ナレッジグラフ構築例

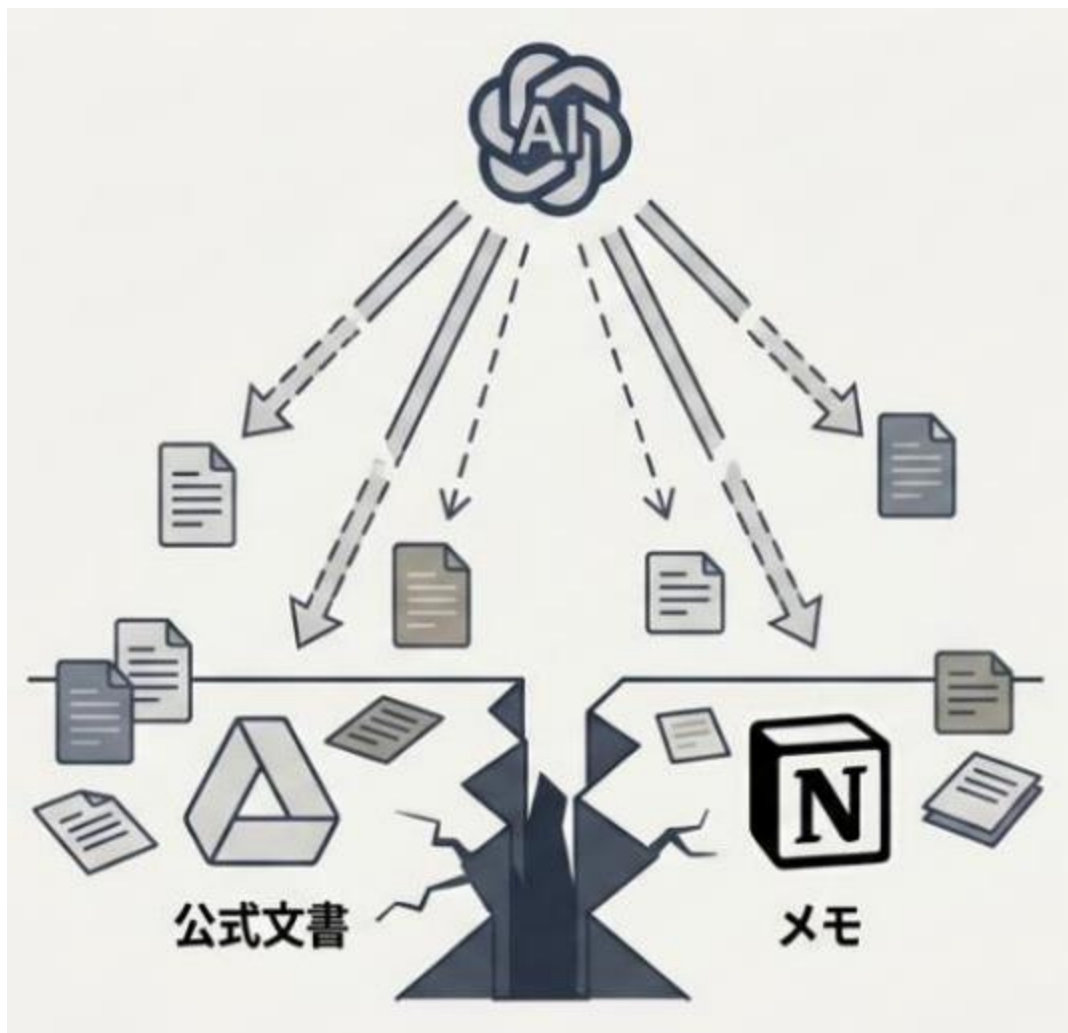
現状の課題と本提案がもたらす変革



なぜ、従来のLLM検索が期待通りに機能しないのか？



情報のサイロ化と「Vector RAG」の限界



用途の混在とサイロ化

Google DriveとNotionが分断され、情報の所在が不明確。

検索精度の頭打ち

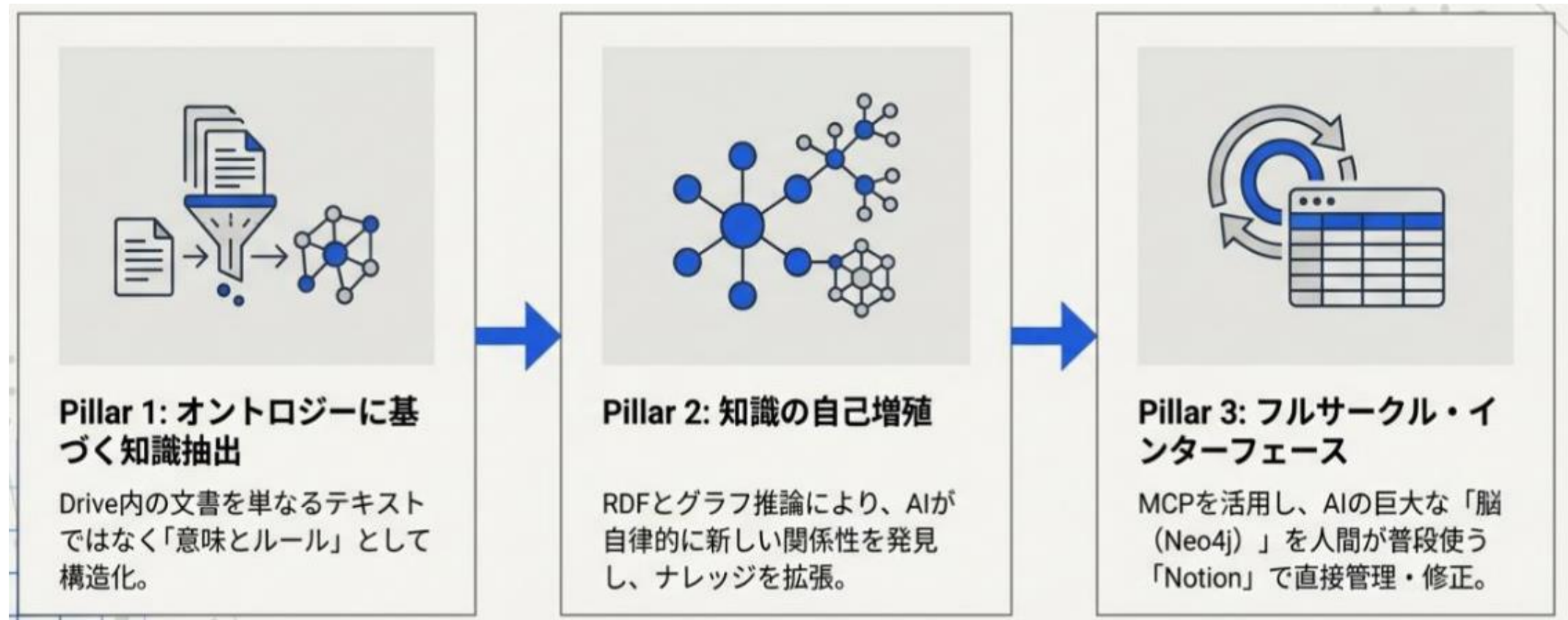
現在のGemini単体（ベクトル検索）は、単語の近さで検索するため、複数文書にまたがる「文脈」や「関係性」を読み解けない。

UIの不在

AIが何を根拠に回答したのか、ユーザー（20名の担当者）が直感的に確認・修正できる場所がない。

「自己成長型」ナレッジ基盤とは？

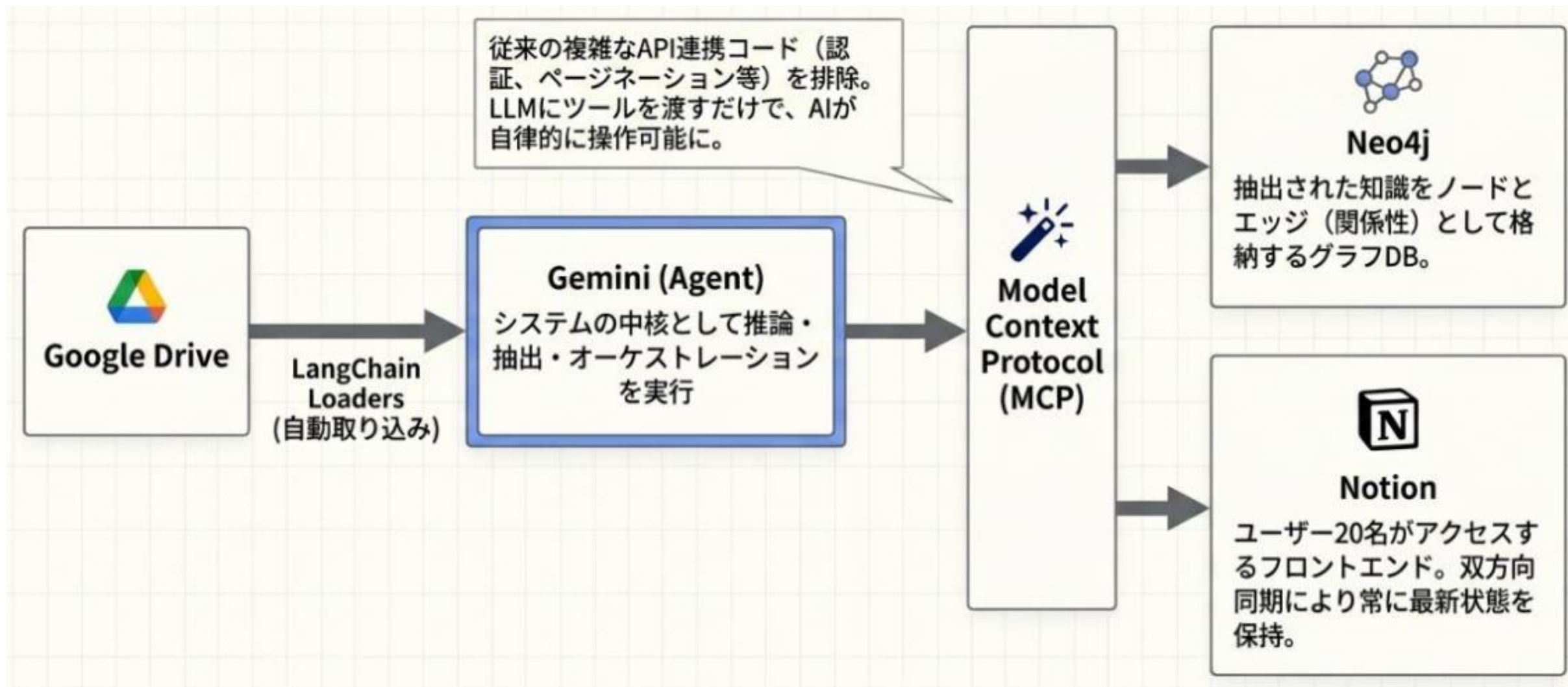
散在するGoogle DriveとNotionの資産をそのまま活かし、Geminiの検索精度を根本から向上させます。



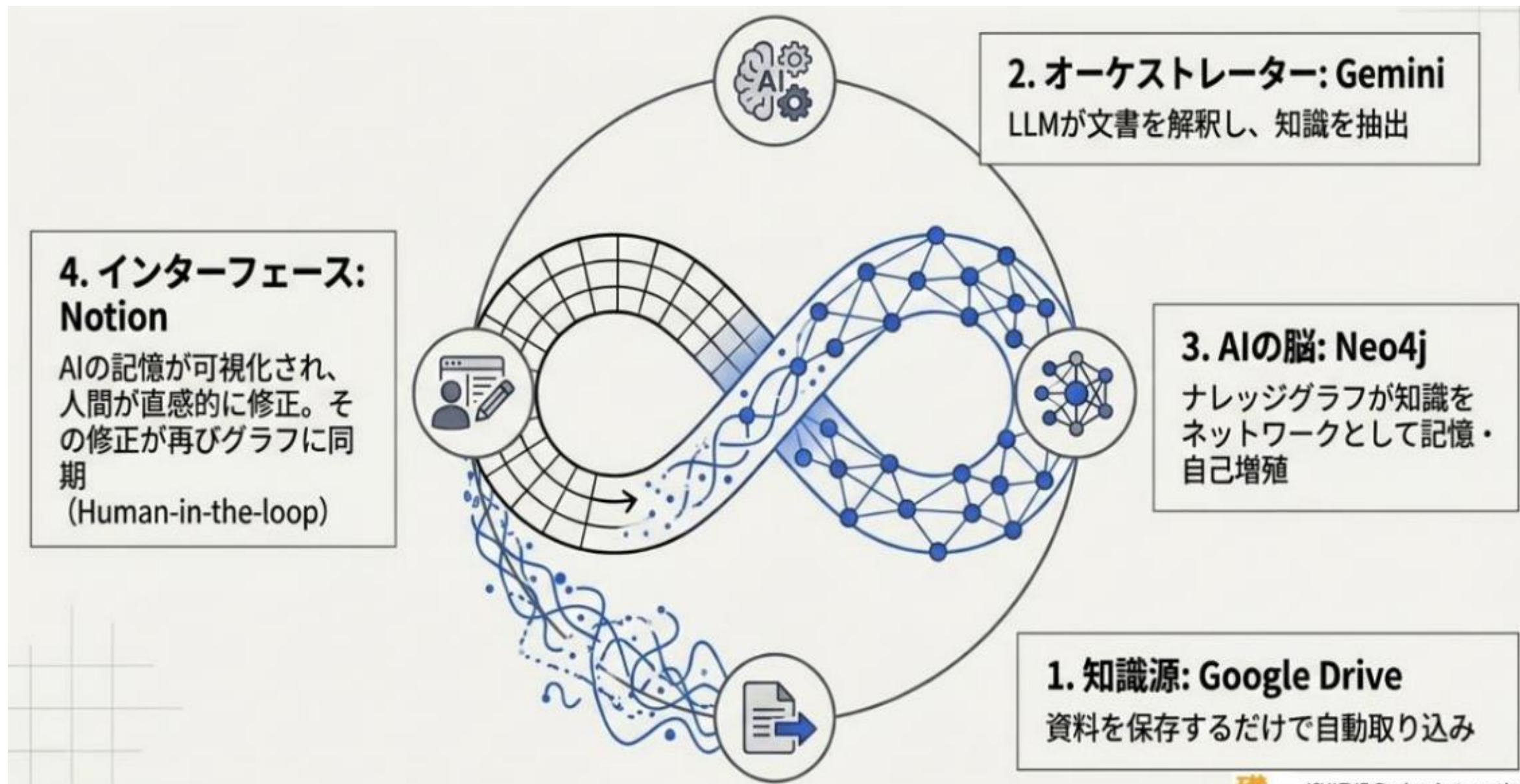
なぜ「グラフ」が必要なのか？

	現在（Vector RAG）	提案（GraphRAG）
情報の表現	埋め込み空間の暗黙的な高次元ベクトル	相互に関連付けられた明示的なエンティティと関係性
文脈の理解	単語の近さに基づく暗黙の文脈（複雑な契約条件に弱い）	定義された関係性による明示的な文脈（条件・例外を正確に把握）
説明可能性	透明性が低く、もっともらしい嘘（幻覚）が発生しやすい	追跡可能な関係性があり、根拠（証拠）に基づく確実な回答
データ処理	表やルールの処理が苦手	構造化・非構造化データの複雑なリンク表現に極めて優れる

システム全体構成図：MCPが実現するシームレスな統合

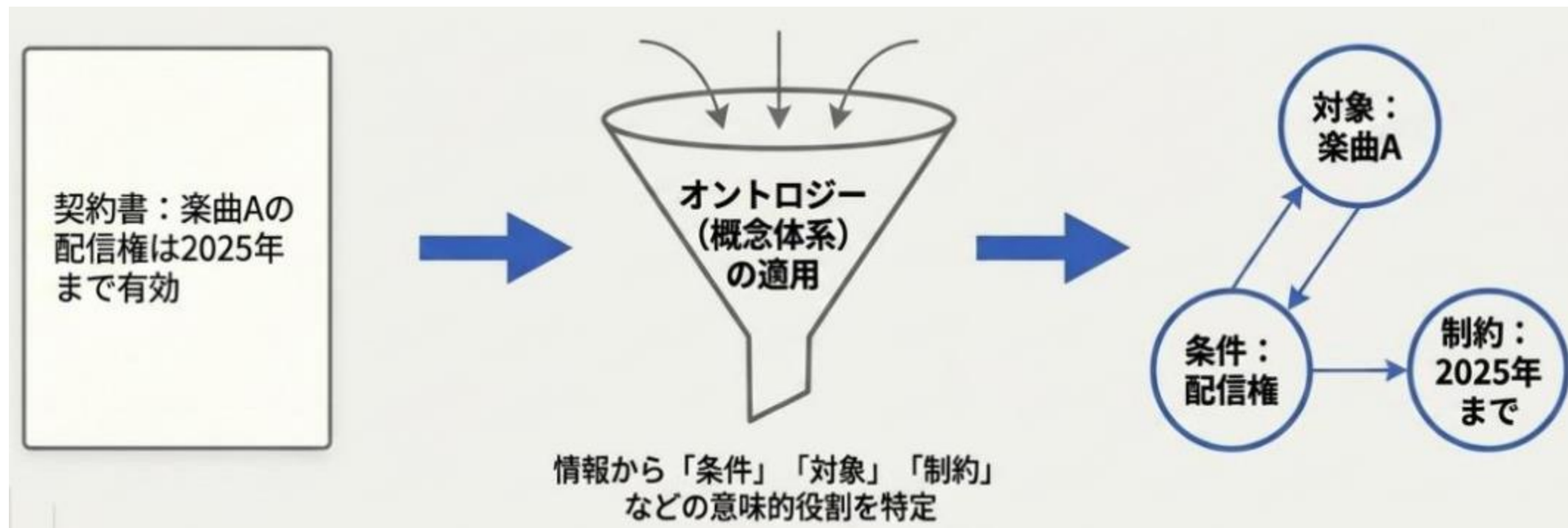


フルサークルアーキテクチャー



オントロジーに基づく知識の構造化抽出

単なるテキストの切り刻み（チャンク化）ではなく、文書を「意味とルール」
として抽出します。

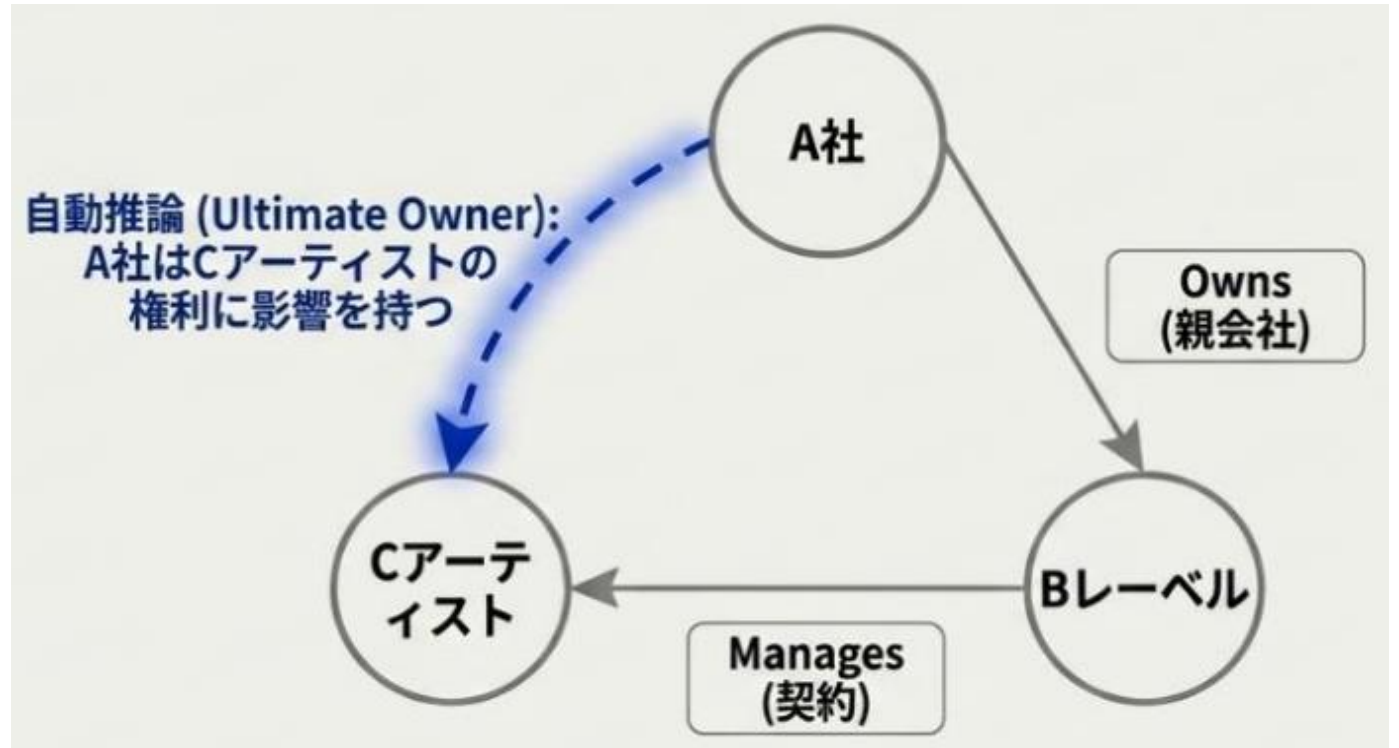


効果: 誰が、いつ、どのような条件で関わるのかという「文脈」を保持したまま知識化するため、
正確な回答の基盤となります。

知識の自己増殖(RDFとグラフ推論)

使えば使うほど、システムが自律的に新しい関係性を推論し、知識ペースを拡張します。

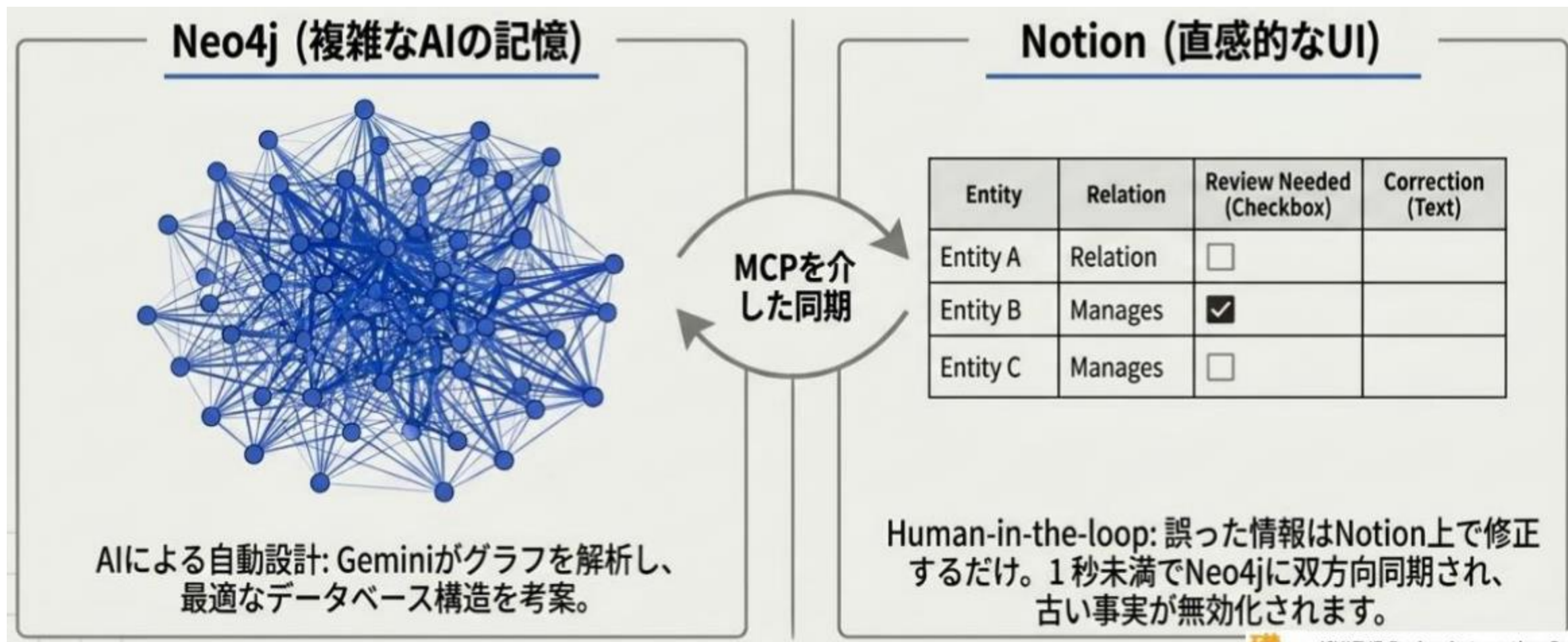
- **RDF (トリプル表現)** : すべての情報を「主語 - 述語 - 目的語」の最小単位で記録。
- **グラフ推論 (Reasoning)** : 既存の2つの事実から、未入力 of 3つ目の事実を自動的に導き出す。



効果: LLMによる新規抽出とグラフ推論が組み合わさることで、手作業による入力なしに、自律的に知識がスケールアップします。

Synapse型 Notionインタフェース

複雑なナレッジグラフを可視化しても人間には扱えません。AIの記憶を、使い慣れたNotionへと変換します。

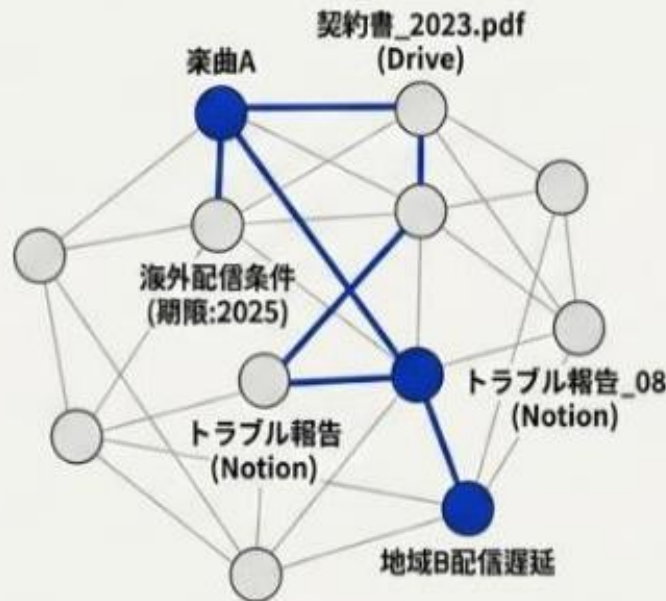


次世代の検索、GraphRAGの威力

User Input

楽曲Aの海外配信に関する契約条件と過去のトラブルは？

メッセージを入力...



文脈を跨いだ検索:

複数の契約書 (Drive) と過去のトラブル対応メモ (Notion) をグラフ上で紐付け。

【回答】

海外配信は2025年まで有効ですが、過去に地域Bで配信遅延のトラブルがありました。

【根拠リンク】

[契約書_2023.pdf \(Drive\)](#)

[トラブル報告_08 \(Notion\)](#)

ブラックボックス化の解消: ナレッジグラフを辿るため、回答の「根拠 (出所)」が常に明確に提示されます。もっともらしい嘘 (幻覚) を防ぎます。



期待される導入効果

情報のサイロ化解消

DriveとNotionが一つの巨大な脳の「知識源」と「インターフェース」として完璧に調和します。

究極の検索体験

20名のメンバー全員が、属人的な記憶に頼ることなく、瞬時に正確な社内規定や過去の経緯にアクセス可能になります。

成長し続ける資産

使えば使うほど推論により知識が増殖し、ユーザーがNotionで修正するたびに賢くなる、真の「自己成長型システム」が実現します。

知識DBの今後の展開

個の属人的知識や職人的ノウハウから、ひとつの「生きた頭脳へ」
ローカルLLMと模倣・強化学習との連携による設計やフィジカルAIへ