Solution Brief

THALES
CYBERSECURITY

RAG（検索拡張生成）AIデータ保護ソリューション

cpl.thalesgroup.com

THALES

CYBERSECURITY

AIの台頭とRAGアプリケーションの登場

人工知能（AI）は、学術的な探求から、世界経済を形作る最も変革的な技術の一つへと急速に進化しました。近年では、GPT、LLaMA、PaLMなどの大規模言語モデル（LLM）の登場により、AIが人間のような文章を生成し、複雑な情報を分析し、さまざまな業界における意思決定を支援するという、驚くべき能力を示しています。膨大なコーパスで訓練されたこれらのモデルは、生産性、顧客エンゲージメント、ナレッジマネジメントの高度化といって分野に飛躍的な進歩をもたらしました。

しかし、その可能性にもかかわらず、LLMは重大な限界に直面しています。訓練データが静的であり、企業独自の知識や特定領域の知識を欠いていることが多いため、出力結果に不正確さや古さ、不完全さが表われることがあるのです。企業にとって、これは大きな障壁となります。組織が必要としているのは、単に言語の流暢さだけでなく、自社データに基づいた、信頼性が高く、文脈を理解し、かつ最新の洞察を提供できるAIシステムなのです。

この課題が、**Retrieval-Augmented Generation**（RAG：検索拡張生成）の台頭を後押ししました。RAGは、外部の信頼できるデータソースを統合することでLLMの能力を拡張する、新しいタイプのAIアプリケーションです。RAGは、企業のナレッジベース、文書リポジトリ、リアルタイムフィード、規制アーカイブにアクセスする動的な検索レイヤーを備えています。モデルの推論に権威ある最新情報を補強することで、RAGは汎用AIと企業向けインテリジェンスのギャップを埋めます。

RAGへの関心の高まりは、AIの潮流が大きく変化していることを示しています。企業はますます、AIの成功を単なる生成能力の高さではなく、結果の正確性、透明性、そして信頼性によって認識るようになっています。金融、医療、政府、重要インフラなどの分野では、誤った回答（いわゆる「ハルシネーション」）や機密データの不適切な取り扱いによるリスクが、自動化のメリットを上回る可能性があります。RAGは、AIの出力を根拠に基づいて生成し、ハルシネーションを低減し、応答の説明可能性と検証可能性を確保することで、こうした課題に直接対処します。

さらに、RAGアプリケーションの拡張性により、企業は従業員、パートナー、顧客が情報とやり取りする方法を刷新できます。従来は手作業の調査や専門知識を必要とした複雑なナレッジ検索が、自然言語クエリを通じて瞬時に提供できるようになります。研究開発の加速からコンプライアンス報告の効率化、顧客サービスの向上に至るまで、RAGは企業全体でのAI導入に向けた実践的な道筋を提供します。

組織がデジタルトランスフォーメーションを加速させる中、RAGベースのシステム導入は急速に拡大すると見込まれています。アナリストは、リアルタイムに企業の知識を統合できるAIシステムが競争優位の基盤となり、データへのアクセス方法だけでなく、戦略的意思決定のあり方そのものを変革すると予測しています。このような状況では、データ保護、コンプライアンス、ガバナンスは不可欠な要素となります。RAGの強力な基盤となるメカニズム、すなわち機密性の高い企業情報を取り込み、保存し、検索し、生成する能力は、同時に新たなリスクも生み出します。強力な保護策がなければ、企業は独自データ、知的財産、個人情報を不正なアクセスにさらす可能性があります。そのため、RAGワークフローをエンドツーエンドで保護することは、単なる選択肢ではなく、責任ある企業AIの基盤となるのです。

RAGワークフローの理解

RAGは、検索メカニズムと生成モデルの2つのコアコンポーネントで構成されています。従来のLLMが訓練データのみに依存するのとは異なり、RAGはクエリに応じて、内部文書、データベース、ライブニュースフィードなど、信頼できる企業固有のナレッジベースから関連情報を検索します。取得した情報を元のクエリと組み合わせることで、RAGは文脈に即した正確な応答を生成します。

ベクトルデータベースはRAGシステムの基盤であり、効率的なセマンティック検索による情報取得を可能にします。典型的なRAGワークフローには次の段階があります。

- **取り込み：**企業は文書（PDF、記事、社内マニュアルなど）をRAGシステムに取り込みます。テキストは「チャンク」と呼ばれる、より小さく扱いやすい単位に分割されます。
- **埋め込み：**埋め込みモデルは、各チャンクをベクトル（意味論的な意味を捉えた数値表現）に変換します。意味が近いチャンクは、高次元空間において互いに近い位置に配置されたベクトルを持ちます。
- **保存：**ベクトルデータベースは、これら数千から数百万のベクトルと、それに対応するテキストチャンクを保存します。
- **検索：**ユーザーがクエリを送信すると、同じ埋め込みモデルを用いてベクトルに変換されます。その後、ベクトルデータベースはk-Nearest Neighbors（KNN）などのアルゴリズムを用いて類似度検索を行い、最も関連性の高い上位k個のベクトルを特定します。
- **生成：**システムは、これら上位ベクトルに紐づくテキストチャンクを取得し、元のクエリとともにLLMに渡します。LLMはそれらを基に、正確で文脈に即した応答を生成します。

タレスのRAGデータ保護ソリューション

タレスは、エンタープライズAIアプリケーションのRAGシステム内にある機密データのライフサイクル全体（取り込み、保存から検索まで）を保護する包括的なRAGデータ保護ソリューションを提供しています。企業は、自社の技術スタック、導入環境、データ分類のニーズ、セキュリティ要件に基づいて適切なソリューションを選択できます。

ユースケース1：取り込み前のデータ検出と保護

CipherTrust Data Discovery and Classification（DDC）ソリューションは、取り込み前に機密データを特定し、分類することを可能にします。さらに、GoogleDataLossPrevention（DLP）、Azure Text Analytics、AWS Comprehendなどのクラウドベースのツールを活用することにより、個々の文書内の機密情報を検出できます。特定された機密データは、企業の情報セキュリティチームが定めたポリシーに従い、タレスの統合データセキュリティプラットフォーム「CipherTrust Data Security Platform」を用いて、トークン化、暗号化、マスキングによって保護することが可能でうす。

ユースケース2：ベクトルデータベースにおけるデータ保護

機密データを取り込み時や埋め込み時に保護できないケースに対して、タレスはベクトルデータベースに保存されたデータを保護するソリューションを提供しています。利用可能な保護手段は、主に次の2つがあります。

- **CipherTrust Transparent Encryption (CTE)**
  CTEは、ベクトルデータベースが使用するストレージ全体を暗号化し、許可されたプロセスのみが暗号化データにアクセスできるようにします。特定のデータベースプロバイダーとの直接的な統合を必要とせず、透過的に動作します。このアプローチは、ベクトルデータベースをSaaSとして利用するのではなく、自社で展開する企業に最適です。
- **CipherTrust Cloud Key Management (CCKM)**
  企業がベクトルデータベースをSaaSとして利用する場合、CTEの適用が困難なケースがあります。そのようなケースでは、CCKMが、SaaSプロバイダーとは分離した形で、暗号鍵を独立して管理することを可能にします。各データベースプロバイダーとの統合が必要となりますが、ほとんどのサービスはCCKMと互換性のあるクラウドベースの鍵管理ソリューション（例：AWS Key Management Service External Key Store）をサポートしています。

ユースケース3：コンプライアンスと脅威検知のためのエンドツーエンドのデータアクティビティ監視

Data Security Fabricの一部であるタレスImperva Data Activity Monitoring（DAM）は、RAGライフサイクル全体にわたるデータベースや非構造化データとのあらゆるインタラクションを、継続的かつリアルタイムで監視します。これには、取り込み、保存、検索、生成に関連するアクセスが含まれます。この機能は、金融や医療などの規制の厳しい業界で特に重要であり、RAGシステムが顧客記録や知的財産などの機密データを扱う場合に不可欠です。

DAMが提供する機能：

- 行動ベースラインの確立：機械学習を用いて、ユーザー、アプリケーション、オペレーションの通常の活動パターン（データ取り込み率、検索クエリ頻度、アクセス行動など）をプロファイリングします。
- 監査およびコンプライアンスレポート：詳細なログにより、誰が、いつ、どこから、どのデータベースやファイルにアクセスしたかを記録します。これらの記録は、カスタマイズ可能なレポートやダッシュボードを通じて、フォレンジック分析やGDPR、SOX、HIPAAなどの規制への準拠に役立ちます。

主な保護機能

- リアルタイム異常検知：内部者による不正利用（例：機密データの過剰な検索）や外部からの探査（例：データ流出の試みを示唆する大量クエリ）などの脅威を特定し、自動アラートとブロックを実行します。
- CipherTrustとの連携によるコンテキスト認識型監視：たとえば、正当な検索時に発生する復号化イベントの検証や、暗号化ストレージへの不正アクセス試行にフラグ付けを行います。
- 脆弱性評価と権限管理：設定ミスの検知、最小権限アクセスの適用、過剰な権限を持つAIサービスアカウントによるリスクの最小化を行います。

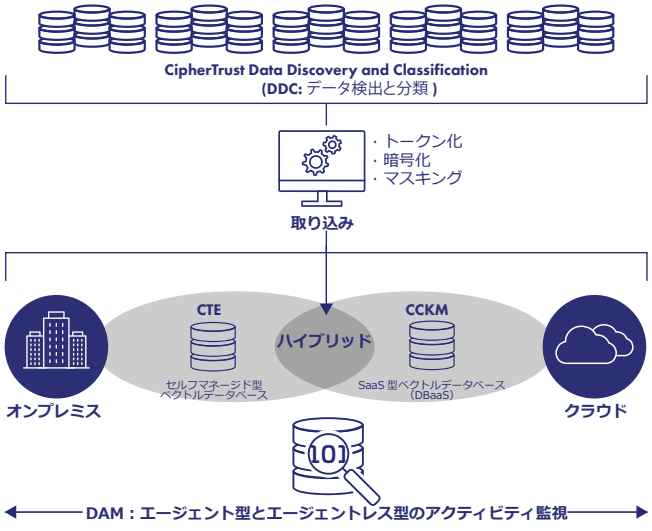
DAMは、エージェント方式とエージェントレス方式の両方に対応し、クラウド、オンプレミス、ハイブリッド環境に導入可能です。また、RAGパイプライン全体のデータフローに対する包括的な可視性を提供しつつ、高スループットのワークロードに対するパフォーマンスへの影響を最小限に抑えます。

エンタープライズRAGの未来を守る

新たなレベルの知性、効率性、競争力を引き出すためにRAG（検索拡張生成）を採用する企業は、複雑なAIワークフローでの機密データの処理に伴うリスクにも対処する必要があります。RAGは、汎用言語モデルとミッションクリティカルなビジネスニーズのギャップを埋める、エンタープライズAIの未来を象徴する存在ですが、その成功は信頼にかかっています。

タレスは、取り込み前の機密データ保護、ベクトルデータベースでの安全な保存、ライフサイクル全体のアクティビティ監視を実現するソリューションを提供することで、企業がセキュリティ、コンプライアンス、レジリエンスを備えたRAGシステムを構築できるよう支援しています。あらゆる段階にセキュリティを組み込むことで、企業は最高水準の精度、コンプライアンス、ガバナンスを満たすRAGアプリケーションを、自信を持って展開できます。

RAGの台頭は単なる技術的マイルストーンではなく、ビジネス変革そのものです。今、安全なRAGを導入する組織は、未来のAI主導型経済―知識が瞬時に得られ、インテリジェンスが文脈に基づき、信頼が最重要となる世界―においてリーダーとなる立場を確立できるでしょう。



お問い合わせ先 – 連絡先情報につきましては、cpl.thalesgroup.com/ja/contact-us をご覧ください。