

データエコシステムが持つ 可能性を完全に引き出す

データレイクエンジン、データ仮想化、
データファブリックの違いを理解する





はじめに

データが世界を動かす現代、企業は膨大な量のデータを生み出しており、それらを効果的に分析すれば貴重な洞察が得られます。しかし、それにはデータ統合、データ管理、データ供給に対応できる堅牢なデータインフラストラクチャが必要です。現在、データにすばやくアクセスするために広く使われている技術として、データレイクエンジン、データ仮想化、データファブリックの3つがあります。データレイクエンジンは便利で拡張性にも優れたSQLクエリ機能を備えているものの、すべての企業のニーズを満たすわけではありません。「A moment of convenience, a lifetime of regret (一時の便利さが一生の後悔に)」と言われるように、将来のニーズを考えずにデータレイクエンジンに投資すると、その後長きにわたって後悔することになりかねません。組織とともに進化できるソリューションを選ぶには、まずデータレイクエンジン、データ仮想化、データファブリックの特徴を比較対照することが重要です。

データレイクの増加

近年、企業の間でデータレイクの導入が広がりを見せています。幅広いソースから得られる構造化および非構造化データを保存できるからです。構造化データしか受け付けず、保存前にデータのクレンジングが必要だった従来のデータウェアハウスとは異なり、データレイクは生データを取り込むことができ、時間と大量のリソースを要する前処理なしで情報を保存、分析できます。

膨らむ一方のデータを前に、企業は行動に結びつく洞察を得るために、データを効率的に処理、分析できるソリューションを必要としています。そうした柔軟性を提供する的就是データレイクです。データレイクは、事前に定義されたスキーマを必要としない、生の形でデータを取り込むことができます。データは、そのまま保存されるため、高速な取り込みが可能です。ACID特性などの機能を追加して進化してきたデータレイクの機能は、データウェアハウスの機能に近づいています。この新たな段階に達したデータレイクアーキテクチャは、しばしば「データレイクハウス」と呼ばれます。

データレイクまたはデータレイクハウスを超える機能が必要な理由

データレイクとデータレイクハウスは、データから価値を引き出すには、まずデータを収集し、物理的に統合する必要があるという原則に基づいて構築されています。多くのユースケースで、これは現実的なアプローチと言えるでしょう。しかし、いくつかの理由から、データが分散されたままになる場合があります。

- **1つのシステムでは対応しきれない場合:** データプラットフォームには、業務用リレーショナルデータベース管理システム (RDBMS)、noSQLソース、データレイク、エンタープライズデータウェアハウス (EDW) など、多くの種類があります。例えば、Amazon Web Services (AWS) では、多様なデータモデルをサポートするために、特定のユースケースに対応した専用エンジンを15種類以上も提供しています。実際、1つのデータプラットフォームで、これらのユースケースのすべてに対応することは不可能なのです。
- **統合が禁止されている場合:** データプライバシーとデータ主権に関するルールは変化しつづけています。一部の国では、データを生成元の国から他国に送信することを制限しています。個人データの第三者への開示を規制する、プライバシー法が施行されている国もあります。これらの国々で活動する企業は、保存または処理を目的としたサードパーティプロバイダーへのデータ送信を、法律で禁止される可能性があります。

最新のデータアーキテクチャは分散化されたデータ環境に対応できなければならない

企業は、これらの一元化されたデータアーキテクチャを元に構築されるデータウェアハウスを、30年以上にわたり変わらず使ってきました。しかし、この間、すべてのデータを単一のプラットフォームに統合できた企業は、皆無とは言わないまでも、ごくわずかです。その原因は、前述したような制約、その他の理由によるものでしょう。これらのプラットフォームの最新のクラウド版は、従来のオンプレミス環境の制約を解消し、スケーリングを簡素化し、最新のデータアーキテクチャにおいて重要な役割を果たしていますが、単一プラットフォームへの統合が非現実的である状況は、今も変わりません。データアーキテクチャの最新化に成功するには、データ戦略に論理データ管理を付け加える必要があります。



足りないコンポーネント：論理データ管理

最新のデータアーキテクチャでは、集中型アーキテクチャで行われる物理的なデータ収集と、分散型アーキテクチャで行われるデータへの論理接続の間でバランスを取る必要があります。論理データ管理では、企業データ資産のすべてについて統一されたデータアクセスレイヤーを設けることで、コピーや複製を行うことなく、あらゆるデータセットに迅速にアクセスできるようにします。論理データ管理には、次のようなメリットがあります。

- **使いやすさ**：利用者は、すべてのデータに1つの場所からアクセスできます。
- **俊敏なデータ統合の選択肢**：1つのプラットフォームで、完全な複製、変換、キャッシュ、リアルタイム連携など、様々なデータ統合の方法を選択できます。
- **一元化されたセキュリティとガバナンス**：アクセス管理やポリシー実装が、1つの論理データアクセスレイヤーに対して適用されるため、一貫性が確保されます。
- **未来への備え**：データへのアクセスが、基盤となるデータソースから切り離されているため、IT部門はビジネスユーザーの業務に影響を与えることなく、基盤のデータストアを自由に変更できます。変更を加えた場合、論理データアクセスレイヤー内のエントリを1つ更新するだけで、データ利用者は変化に気付くことなく、新しい場所へと自動的に誘導されます。ビジネスユーザーが必要なスピードで業務を続ける一方で、IT部門も自分たちのスピードで変更を加えることができるのです。

進化を続ける論理データ管理エコシステム

論理データ管理という分野は決して新しいものではありません。多くの確立された技術と新しい製品カテゴリーが存在します。このホワイトペーパーでは、分散データの統合と管理に使用される、最も一般的な3つの手法について、それぞれの強みと制約を明らかにし、読者が情報に基づいた判断を下せるようにします。

- **データレイクエンジン**では、データレイクに保存されたデータに対してSQLクエリを実行できます。重要な特徴の一つが、データレイクに保存された大規模で複雑なデータセットに対する対話型SQLクエリの実行性能（速度）が改善される点です。このため、より低速のバッチ処理を重視していたHadoopベースの第1世代データレイクに対比させて、これらのシステムを「クエリアクセラレーター」と呼ぶアナリストもいます。通常、これらのシステムでは、従来型の超並列処理（MPP）データベース管理システムと同様のクエリアーキテクチャを使用し、多くの場合、運用とメンテナンスに専門のスキルが必要です。データレイクエンジンの中には、外部ソースへのリアルタイムアクセスが可能なものもありますが、その機能は限られている場合がほとんどです。どのような機能を提供され、どのような制約が存在するのかを把握することが重要です。
- **データ仮想化**により、統合された論理ビューを介して複数の異なるソースからデータにアクセスできるようになるため、ユーザーに対して、データソースの根底にある複雑さを意識せずに、シンプルで標準化された形でデータを表示できます。したがって、異なるフォーマット、構造、場所に保存されているデータであっても、1つの統合されたデータソース内にあるかのようにクエリや分析を実行できます。データ仮想化では、企業全体でセマンティックデータモデルを使用することで、データアナリスト、開発者、ビジネスユーザーが最小限のトレーニングで、簡単にデータを理解し利用できるようにします。
- **データファブリック**は、ハイブリッド/マルチクラウド環境全体で、データを統合し、データエンジニアの作業を自動化します。データファブリックは、データソースやデータタイプが異なっても効率的で自動化されたデータ管理が可能な、包括的な自動化機能を提供します。これらの機能には、データ統合、データの前処理と実処理、データのカatalog作成と探索、データのガバナンスとセキュリティなどがあります。データファブリック構想の中核をなすのは、多くの場合、データ仮想化であることに注意してください。データ仮想化により、毎回、環境の異なるポイント間で物理データを複製するのではなく、環境にまたがる論理データアクセスが可能になります。しかし、データ仮想化からデータファブリックに移行するには、他にも多くの機能が必要です。



このホワイトペーパーを最後まで読めば、上に述べた重要な技術の内容、メリット、注意点をより深く理解できるでしょう。本書の目標は、データエコシステムの価値を最大化し、データ主導の最新の環境で大きな成功を手にするための、最適な選択を行えるようにすることです。おわかりのように、これから説明する3つの技術は、革新的な機能の集まりです。まず、3つの技術に共通する機能について見ていきましょう。

データ管理への論理的アプローチに必要な機能

データファブリックのアーキテクチャは、従来のデータレイクアーキテクチャに比べ、データ管理の方法が包括的で柔軟であることから、支持を広げてきました。複数のソースからのデータを統一し、高度なデータ処理および分析機能を提供することで、より深い洞察をもたらし、よりの確な意思決定を可能にします。これに加えて、より包括的なガバナンスとセキュリティ管理も実現するため、データプライバシーの確保や規制への準拠が容易になります。

テクノロジー業界にはよくあることですが、新しいアーキテクチャに人気が出ると、その領域のサービス提供において有利な地位を確保しようと、多くのベンダーが先を争って参入してきます。しかし、すべてのベンダーが、必要な機能に等しく対応しているわけではありません。このセクションでは、多くの業界アナリストたちによって確立された基準に基づいて、様々な製品カテゴリーの主要機能を分類したうえで、それぞれについて解説します。このセクションの目標は、以下の通りです。

- 利用できる機能の把握
- 読者自身の現在および将来にわたる具体的ニーズの検討
- 読者自身の評価基準の定義

これによって、様々な製品カテゴリーの各種ソフトウェア技術を効果的に評価し、ビジネスニーズに最適な製品を情報に基づいて判断できるようになります。

注:本書末尾の「推奨参考文献」のセクションには、下表をまとめる際に役に立ったアナリストのレポートへのリンクが収録されています。



		データレイクエンジン	従来のデータ仮想化	DENODO PLATFORM	データファブリック
1	データアクセスレイヤー	●	●	●	●
2	データレイククエリの最適化	●		●	●
3	フェデレーテッド/ 分散クエリエンジン	●	●	●	●
4	データ複製 (永続的)	●	●	●	●
5	メタデータリポジトリ	●	●	●	●
6	セキュリティおよび ガバナンスレイヤー	●	●	●	●
7	利用レイヤー	●	●	●	●
8	ユニバーサル セマンティックレイヤー	●	●	●	●
9	設計インターフェイスとツール	●	●	●	●
10	拡張データカタログ		●	●	●
11	アクティブメタデータ			●	●
12	レコメンデーションエンジン			●	●
13	データの前処理および供給レイヤー			●	●
14	オーケストレーションおよび DataOps			●	●

最小限の対応
 最大限の対応

- データアクセスレイヤー**は、複数のソースに含まれるデータのアクセスと統合を簡素化します。この際、ソースがSQLをネイティブにサポートしていない場合でも、通常はSQLを使用します。データ利用者には、基盤となるデータソースの場所やフォーマットに関係なく、一貫性のある統一されたデータビューを提供します。データアクセスレイヤーには、データベース、アプリケーション、API、Webサービスなど様々なデータソースに接続するコネクタまたはアダプターが含まれます。各種技術を評価する場合は、利用できるコネクタ、それらのプッシュダウン機能の複雑性、対応しているシステムの種類の検討する必要があります。現在または将来使用する可能性があるデータソースの中で、一つでも対応していないものがあるソリューションに対しては、投資を控えることをお勧めします。
- データレイククエリの最適化**とは、データレイクに対して実行されるクエリを最適化することです。これによって、クエリが許容できる性能レベルで効率的に実行されるようにします。そのための技術として、超並列処理 (MPP) や様々な特徴 (フレーバー) を持ったクエリアクセラレーターがあります。データレイククエリの最適化は、これらのツールを選定する際に考慮すべき重要な機能です。クエリの性能が強化され、データ転送のコスト削減やシステム全体の効率向上につながるからです。通常、データレイクエンジンは、情報の主なソースがデータレイクである場合に、最大効率で動作するように設計されており、分散環境では性能が劣化することに注意してください。データのすべてまたは大部分がデータレイクに保存されるかどうか、そうでない場合は、性能にどのような影響があるのかを理解しておく必要があります。

3. **フェデレーテッド/分散クエリエンジンを使用すると**、複数のソースからのデータに対して、あたかもそれらが単一のデータベースに保存されているかのようにアクセスし、クエリを実行することができます。この種のクエリエンジンは、複数のデータソースにわたるクエリを実行し、その結果を1つに結合してユーザーに戻します。データベース、アプリケーション、Webサービス、データレイクなど、異なる種類のソースからのデータにアクセスし、統合できます。関連するデータソースの特定、必要なデータの取得、単一ビューへの結果の結合までがフェデレーテッドクエリエンジンの役割です。フェデレーテッドクエリエンジンは、スタンドアロンのクエリアクセラレーターに使用されている数々の手法を実行するだけでなく、高度なクエリプッシュダウン、クエリリライト、クエリアクセラレーション、クエリキャッシングなど、複数のデータソースにまたがり実行されるように設計および最適化されたその他の手法も提供します。
4. **データ複製(永続的)**とは、データのコピーを作成、管理することです。データ複製には、様々な形態があります。抽出-変換-読み込み(ETL)プロセスは、ポイント間複製に使われます。抽出-読み込み-変換(ELT)プロセスは、データレイクまたはデータレイクハウス内の動作、外部APIからのマイクロバッチ処理、ストリームチャネルからの取り込み、キャッシングおよびアクセラレーション構造に使われます。ユースケースはそれぞれ異なり、最新のアーキテクチャでは、こうした複製作業の作成、管理、運用に利用できる手法を、柔軟に選べなければなりません。
5. **メタデータリポジトリ**は、スキーマ、データの種類、データの統計情報、データの関係性など、データソースに関するメタデータが一元的に保存されるコンポーネントです。ソースに対するマッピング、リネージ、変換方法など、利用者のデータモデルに関する情報も提供します。メタデータリポジトリには、アクティビティに関するメタデータも保存できます(後述の「アクティブメタデータ」を参照)。メタデータリポジトリは、クエリ最適化の基盤となるだけでなく、変更管理、セキュリティ設定、セルフサービス機能の一部(他のセクションで詳述)など、データガバナンスの基盤にもなります。
6. **セキュリティおよびガバナンスレイヤー**は、特定のデータに対する認証やアクセスを管理する高度なメカニズムを提供します。例えば、個人特定情報(PII)としてマークされた列に対しては、情報を共有しない権限設定、または許可されたユーザーの一部しか共有できない権限設定が可能です。これは、細部が重要な領域です。検討中のソリューションが、属性に基づくアクセス管理(ABAC)システムなど、データアクセスに関するセキュリティポリシーを適用するための高度な機能に対応しているかどうかを把握する必要があります。ポリシーエンジンは、セキュリティポリシーを定義して適用します。ポリシーはデータの機密性、ユーザーロール、ビジネス要件など様々な要因に基づいて設定できます。検討中のツールが、既存のセキュリティおよびガバナンスツールと連携するかどうかを確認しておくことをお勧めします。
7. **利用レイヤー**を介して、データ利用者は、すべてのデータに一元化された方法でアクセスできます。このレイヤーは、エンドユーザーが自分の望む形態でデータを受け取れるようにします。ほとんどのシステムでSQLクエリにJDBCまたはODBCを使用できますが、データを利用する可能性があるユーザーを漏れなく考慮し、他に必要なオプションがないかを検討する必要があります。その他のオプションとして、ADO.NET、MDX、RESTful Webサービス、OData、GraphQL、GeoJSONに加え、Microsoft Excel/SQLへのエクスポート、Kafkaのサブスクリプション、JMSメッセージキューなどがあります。
8. **ユニバーサルセマンティックレイヤー**を使用することで、エンドユーザーが簡単に理解し、利用できるデータ構造を定義できます。セマンティックモデリングに関連する機能は、わかりやすい名前と説明が付けられた派生データモデルの定義、社内指標の定義、タグ付けや分類、ガバナンス管理(推奨または非推奨など)、データ品質ラベルなど、多岐にわたります。これらはいずれも、データをわかりやすく使いやすきものにすることで、データアクセスを民主化するという目標を共有しています。この目標を重視するならば、セマンティックスは、間違いなく重要な評価基準になるでしょう。
9. **設計インターフェイスと関連するツール**は、システム管理を支援します。その機能は、コマンドラインのみのツールから、データエンジニアやエンドユーザーが簡単にシステムと対話できるWeb UIまで、ベンダーごとに大きく異なります。どのような種類のインターフェイスが快適に利用できるかを検討してください。ただし、組織の全ユーザーが利用できるようにデータの民主化を目指している場合は、想定されるあらゆるユーザーが直感的に使用できるツールであることを確認する必要があります。
10. **拡張データカタログ**は、組織全体のデータ資産に関する検索可能な一元的なインベントリと、メタデータ、データリネージ、その他のデータに関する情報を提供します。拡張データカタログは、データ資産をその内容と構造に基づいて自動的に分類し、タグを付けます。データスチュワードやアナリストがデータ資産に手でタグを付けたり、注釈を付けたりして、ビジネス定義、データ品質指標、利用状況などの情報を補足できるようにする機能もあります。拡張データカタログによって、ユーザーは組織全体にわたるデータ資産を迅速かつ簡単に探索できるため、コラボレーションや意思決定が強化されます。

11. **アクティブメタデータ**は、データエンジニアの作業を自動化して、複雑性を軽減し、ビジネスの成果と価値を最適化します。従来、「メタデータ」という言葉は、他のデータを説明するデータを意味していました。しかし最新のデータ管理では、アクティブメタデータを作成するために使用されるトランザクションログ、ユーザーログイン、クエリ最適化計画などの要素を含む、技術的なメタデータも導入されています。アクティブメタデータは、システムの性能やユーザーとシステムとの相互運用性などの指標を記述するデータです。組織のデータインフラストラクチャやコストの責任分担を促進するFinOpsの取り組みにも使用されています。
12. **レコメンデーションエンジン**は、機械学習アルゴリズムによって、検索クエリ、データアクセス、データダウンロードなどのアクティブメタデータを分析し、それらのパターンを特定することで、システムの運用効率を高める推奨事項（レコメンデーション）を提案します。ユーザーが関心を持つ可能性がある他のデータ資産についても提案します。提案に従えば、人手によるデータ管理業務が減り、ビジネスユーザーは必要なデータをより迅速に入手できるようになります。
13. **データの前処理と供給レイヤー**は、データを探索、変換して新しいデータセットを作成できるサンドボックスのような環境内で、共通セマンティックレイヤーを使用できるようにします。これは、セルフサービスを実現したり、利用者が最後の変換を実行してデータセットを最終的な消費者形態に変更したり、利用者からのフィードバックの収集やファブリックの改善のクラウドソーシングを可能にしたりするための重要な機能です。
14. **オーケストレーションおよびDataOps**は、データパイプラインとワークフローを管理して、データ関連のプロセスを自動化します。オーケストレーションとは、複数のシステムやプラットフォームにまたがることの多い、複数のデータ関連の作業やワークフローを協調させることを意味します。複数のソースからのデータを統合するデータパイプラインの管理や、データの処理や分析の作業の調整を伴う場合があります。これに対してDataOpsは、DevOpsの原則や実践方法を取り入れたデータ管理の手法です。DataOpsでは、コラボレーションと自動化、データ関連のプロセスのスピード、品質、信頼性の改善に重点が置かれます。

データレイクエンジンは当初、データエンジニアが一元化されたりポジトリ内の構造化および非構造化データを大量に処理できるように設計され、その後、外部のデータソースにも対応できるように強化されました。これに対して、データ仮想化は、任意のソースからのデータに、あたかも単一のデータベースに保存されたデータであるかのようにアクセスし、クエリを実行できるようにすることを目的に設計されました。データレイクエンジンのクエリは、すべてのデータがデータレイク内に存在する場合であれば極めて効率的であるものの、外部エンジン（自社のEDWなど）または分散データソースに散在するデータへのクエリの場合は、一般的にデータ仮想化ソリューションの方が高い性能を発揮します。

ユーザーエクスペリエンスも、2つの製品カテゴリーで異なる傾向があります。データレイクエンジンは、コマンドラインが当たり前のインターフェイスに慣れた技術系ユーザーを対象にしていたのに対し、データ仮想化では常に、ビジネスユーザーが必要なデータを見つけてアクセスできるように支援することを目指してきました。このため、通常、データ仮想化のツールの方が、ビジネスフレンドリーなインターフェイスを備えています。

データファブリックもあります。データファブリックは上記の両製品カテゴリーが最も強みとする機能を駆使し、ガバナンス、セキュリティ、データ統合などの高度な機能を含む、複数のソースやシステムにまたがる統合されたデータビューを提供します。データファブリックと、本稿で取り上げた他の製品カテゴリーのもう一つの違いは、従来手動で行ってきた作業の自動化に重点が置かれていることです。

データファブリックは、上記の機能をすべて提供することで、現代のデータ主導の企業が求める要件を満たすための適応と進化が可能な、最新のデータアーキテクチャの基盤を確立します。アジリティ、柔軟性、イノベーションを推進し、データレイクはもとより企業内のあらゆるデータ資産に秘められた潜在能力を解き放ち、企業の成長につながる洞察をもたらします。

Denodoは、シンプルな実装を目指す中小企業から、数千人ものユーザーが分散データエコシステムに対して信頼性の高いクエリを毎日数百万回も実行する世界有数の大手企業まで、幅広いお客様とお取引いただいています。

次のセクションでは、世界29か国で約180万台の自動車を管理している、CaaS（サービスとしての自動車）業界の世界的リーダーであるLeasePlanの事例を紹介します。

Denodo Platformを使用して データファブリックを構築したLeasePlan

LeasePlanは、従来のアナログなビジネスモデルを完全にデジタル化したいと考えていました。サービスポートフォリオを拡大し、革新的な製品とサービスによって収益を生み出す事業を推進することで、完全デジタル対応のCaaS企業に生まれ変わることを目指したのです。LeasePlanの課題は、行動データ、マーケティングデータ、交通情報、ソーシャルメディア情報、サービスとメンテナンスのデータなど、膨大な量の情報の維持管理でした。しかも、これらのデータはサイロ化された異種混在のデータソース（SAP、Salesforce、IBM DB2、Snowflakeなど）内に分散していました。ビジネス上の的確な意思決定、プロセスの最適化、新しいビジネスモデルの創生、EU規制への準拠を可能とするには、データ統合とデータ供給に問題がありました。



そこでLeasePlanは、データ仮想化を駆使した論理データ管理プラットフォームDenodo Platformを導入しました。このプラットフォームにより、LeasePlanは最新のデータソースからクラウド、レガシーのデータソースまで、主要データソースのすべてをカバーし、一元的にアクセスできる論理データファブリックを確立できました。論理データファブリックは、抽象化によって直接アクセスの複雑性を解消する、データ仮想化を使用します。この機能を使って、LeasePlanは多岐にわたる新しいデータソースや大量のデータを処理するアプリケーションへの接続も実現しました。その結果、業務に影響を与えずに、必要に応じてデータ技術を簡単に切り換えることができました。LeasePlanの論理データファブリックは、導入後すぐにレガシーデータウェアハウスに代わり、一元データアクセスレイヤーとして機能しはじめました。論理データファブリックは、一元化されたデータガバナンスフレームワークを備えることで、LeasePlanの顧客を保護し、最高レベルのデータ安全性を確保します。

論理データファブリックの実装によって、車両管理者に的確な洞察がもたらされ、これがカスタマーエクスペリエンスの強化につながりました。ある1台の車両の情報を、他の車両の履歴情報と結合することで、車両の整備が必要となるタイミングの予測精度が、これまでよりはるかに向上し、不具合が発生する前に顧客を希望する修理工場に案内できるようになりました。顧客の車両整備予約のスケジュールリングまで実現しました。今では、BIレポート用に、オンプレミスのデータをダウンタイムなしでクラウドに移行できます。さらに、世界各国の様々な事業部門で、レガシーデータウェアハウスのデータマートをクラウドに効果的にオフロードしました。

LeasePlan Digitalのエンタープライズデータおよびアナリティクス担当ディレクターSumit Arya氏は、次のように述べています。「Denodo Platform上に構築された新しい論理データファブリックは、データのセルフサービスの基盤として、LeasePlanのグローバルデータハブになくてはならないコンポーネントです。新しいビジネスサービスが生み出されるとともに、カスタマージャーニーのあらゆる段階でドライバーを的確にサポートできるようになりました」



「Denodo Platform上に構築された新しい論理データファブリックは、データのセルフサービスの基盤として、LeasePlanのグローバルデータハブになくてはならないコンポーネントです。」

Sumit Arya氏、LeasePlan Digitalのエンタープライズデータおよびアナリティクス担当ディレクター

Denodo Platformについて

Denodo Platformは、最新のデータファブリック戦略の基盤を確立する、論理データ管理、データ統合、データ供給のためのソリューションです。お客様は、Denodo Platformを使って、セルフサービスのBI、高度な分析、ハイブリッド/マルチクラウド統合、エンタープライズデータサービスのソリューションを実現しています。Denodo Platformは、業界アナリストが必要性を認めた機能と本稿で取り上げた機能のすべてに、高い信頼性で対応します。これによって、企業は異種混在データソースのすべてに対して、シームレスに統合されたアクセスが可能な、論理データファブリックを確立できます。

Denodo Platformには、論理データファブリック構築において業界をリードするソリューションとして、他にも次のような機能と特徴を備えています。



統合データカタログ:

自動レコメンデーション、コラボレーション強化、プロファイリング、スマート検索機能などによってデータアクセスを合理化します。



普遍的な接続性:極めて広範

なデータソースに簡単に接続できます。サードパーティのシステムに対する200以上のコネクタが用意されています。



組み込みMPPデータレイ

クエンジン:データ利用者が、Prestoに基づく拡張性に優れた高性能のSQLエンジンを利用してデータレイク内のデータにアクセスできるようにするとともに、直感的なGUIによりデータアクセスを簡素化します。



スマートクエリアクセラレー

ション:インテリジェントなキャッシュ、コストに基づく最適化、AIによるレコメンデーションエンジンによって、複数ソースに対するクエリ実行を最適化します。



属性に基づくアクセス管理

(ABAC):セマンティックレイヤー、タグ付け、セッション属性を組み合わせることで、データの機密性、ユーザーロール、ビジネス要件などの様々な要因に基づいてグローバルに定義される、きめ細かいアクセスポリシーを使用できます。



FinOps:運用コストを可視化

して、クラウドインフラストラクチャのコスト管理を最適化します。



データ供給の幅広い選択肢:

ビジネスユーザーがSQLクエリにJDBCまたはODBC以上のものを使用できるだけでなく、開発者が、RESTやGraphQLのAPI、Kafka、JMSメッセージキューをはじめとする多彩なデータ供給方法を選択できます。



Denodoは、2022年第2四半期のForrester Wave: Enterprise Data Fabric、2022年のGartner Magic Quadrant for Data Integration Toolsで、業界アナリストによってリーダーとして認められました。

お客様も、この評価に賛同しています。2023年Gartner® Peer Insights™ “Voice of the Customer”: Data Integration Toolsレポートで、カスタマーズチョイスに選出されました。3年連続で、この賞を授与されたデータ管理ベンダーはDenodoだけです。

今すぐ始める

このホワイトペーパーによって、論理データファブリックのメリットを明らかにできたものと思います。論理データファブリックは、クラウドの柔軟性と拡張性を1つの場所から使いやすい形で提供します。ユーザーは、データを安全に検索、アクセス、統合、共有することで、ビジネス上の意思決定をより多くの情報に基づいてタイムリーに下すことができます。

結論

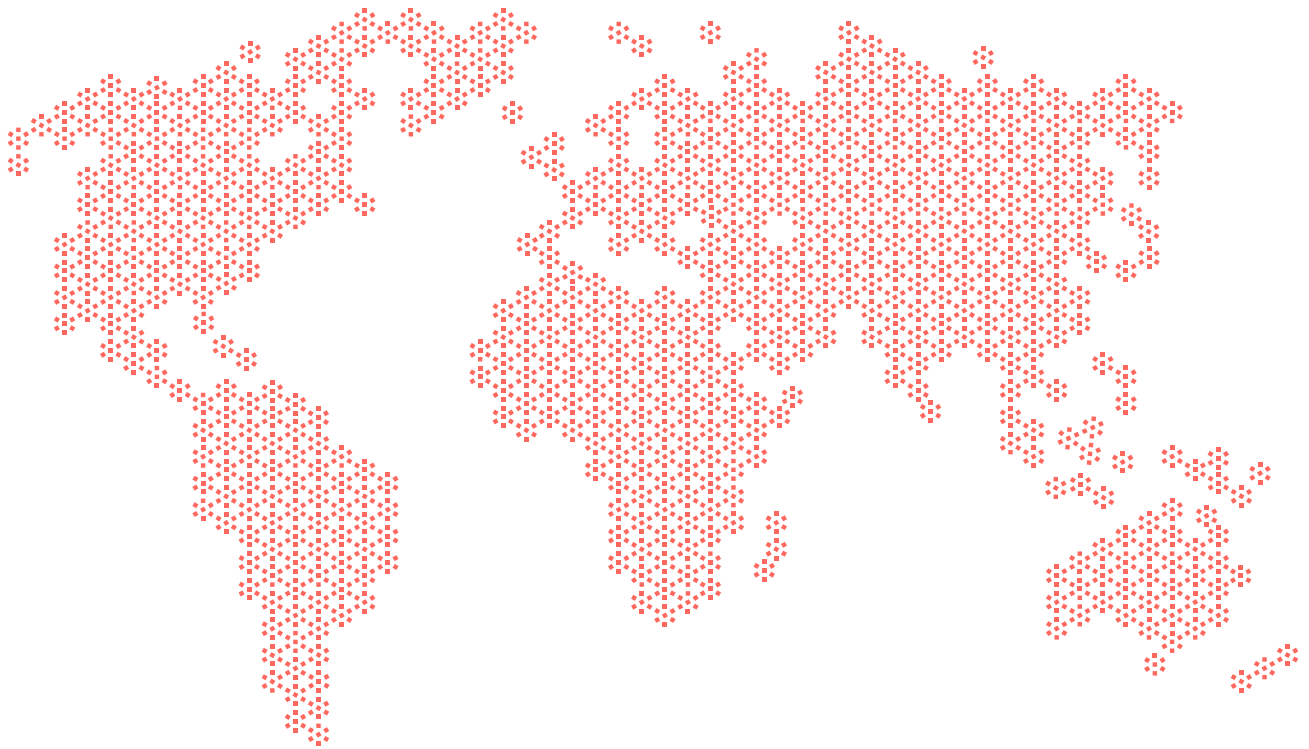
データは分散するのが常です。しかし、論理データファブリックなら、データを一元リポジトリに物理的に移動することなく、必要なときにリアルタイムでデータに接続する手段を提供できます。

このホワイトペーパーでは、業界アナリストが必要性を認めた主要機能の中から、論理データ管理ソリューションの広範な機能を14件紹介しました。現在市販されているソリューションは、それぞれが若干異なる割合で、これらの機能を備えています。特異なニーズを持ち、かつ今後ビジネスが大きく変化する可能性がない企業であれば、データレイクエンジンまたは、自社のニーズに特化した従来のデータ仮想化ツールで十分対応していけるかも知れません。しかし、柔軟性を最大化し、組織の投資を将来にわたって有効なものにするには、論理データファブリックの形で上記の14の機能のすべてに対応するソリューションの導入を検討する必要があります。そうしたソリューションであれば、現在および将来のニーズに合った選択肢のほとんどを提供し、最新化されたデータインフラストラクチャのメリットを余すところなく引き出すことができます。現在の極めて競争が激しく、目まぐるしく変化するビジネス環境では、そうしたソリューション導入の成否が、今後数年間における企業の成功と低迷の分かれ道になるかも知れません。

推奨参考文献：

1. Gartner - Demystifying the Data Fabric (ガートナー：データファブリックの謎を解明する)
2. Gartner - Assessing the Relevance of Data Virtualization in Modern Data Architectures (ガートナー：最新データアーキテクチャにおけるデータ仮想化の適用可能性を探る)
3. The Forrester Tech Tide™: Enterprise Data Integration, Q4 2021 (The Forrester Tech Tide™：2021年第4四半期エンタープライズデータ統合)
4. Denodo - Logical Data Fabric Whitepaper (Denodoホワイトペーパー：論理データファブリック)





Denodoはデータ管理のリーダーです。受賞歴のあるDenodo Platformは、セルフサービス型BI、データサイエンス、ハイブリッド/マルチクラウド・データ統合、およびエンタープライズデータサービスを可能にする論理的アプローチを使用した、業界をリードするデータ統合、管理、および配信プラットフォームです。400%以上のROIと数百万ドルの利益を実現し、30を超える業界の大企業やミドルマーケット企業のDenodoの顧客は、6か月足らずでの投資回収を実現しています。

Webサイト www.denodo.com/ja | Email info@denodo.com | Discover community.denodo.com

