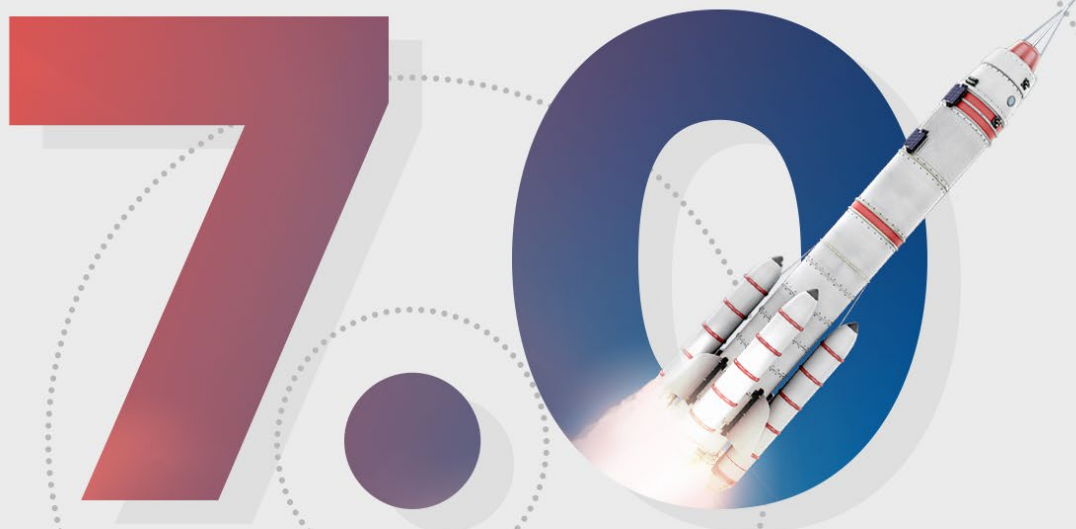


# Denodo Platform 7.0 - 新機能





## Denodo Platform 7.0 - 新機能

今日データ管理はますます複雑になり、ビッグデータシステムはますます重要になっています。また、企業は、かつてないほど多くの処理をクラウドに移行しています。また、ビジネスユーザは情報にセルフサービスで簡単にアクセスできる必要があり、IT部門によりビジネスの需要をすべて満たすのが困難になっています。

今回新たにリリースされるDenodo Platform 7.0は、今日企業が抱えている複雑なデータ管理の問題を緩和し、データ管理を高いレベルで行えるようにするためのソリューションです。Denodoでは、データ仮想化による統合データ配信レイヤーを提供するビジョンのもとに、ビジネスユーザとアプリケーションのデータに対して単一のソリューションで対応できるようにしています。これにより、ユーザが複雑な環境に直面することなく、ビジネスを行う上でデータを問題なく制御できるようにすることで、実際のITインフラストラクチャとのギャップを埋めることができます。

Denodo Platform 7.0の新しいリリースは、ビジョンを具現化するために、ビジネスユーザが必要な情報を簡単にブラウズし、検索できる機能を提供するだけでなく、IT部門によるセキュリティやガバナンス統制を効かせられるようにします。

オンプレミスとクラウドの両方にあるさまざまなデータ保存先の間において、データストア、データボリューム、データセットの配信が異種混在しているため、データ管理はますます複雑になっています。Denodo Platform 7.0には、刷新されたパフォーマンスの機能があり、これにより非常に高い需要であるビッグデータやアナリティクスの環境にも対応できるように、かつてなかった拡張性が実現されています。今日のハイブリッドマルチクラウド環境に対応するために、Denodo Platform 7.0は、コンテナリゼーションを通じて、開発運用、クラウドのデータ仮想化、クラウドアプリケーションのデプロイメントを自動化する新しい機能を導入しています。

Denodo Platform 7.0の新機能を、次の主要な特徴に従って分類し、確認してみましょう。

- 比類なきパフォーマンス:インメモリ並列処理と動的クエリ最適化を融合したパワー
- 動的データカタログ
- ITペルソナとビジネスペルソナ
- マルチクラウド、ハイブリッド、エッジによる、柔軟性の最大化
- 自動ライフサイクル管理

5つの主要特長は、ITユーザにとってはすべてが、ビジネスユーザにとっては2番目、3番目、4番目が重要です。

## DENODO PLATFORM 7.0

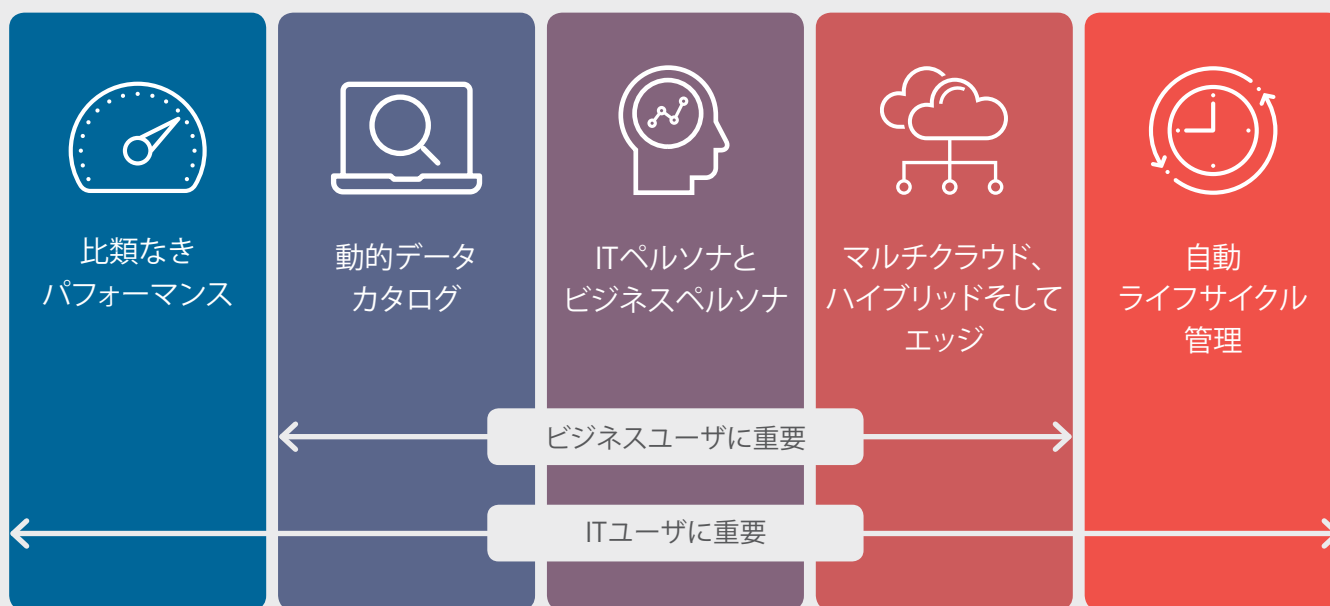


図1:Denodo 7.0の主要特長

データサービス公開、データガバナンス、ならびにセキュリティー、接続性、データ変換の新機能についての情報も追加しています。

# 比類なきパフォーマンス:インメモリ並列処理と動的クエリ最適化を融合したパワー

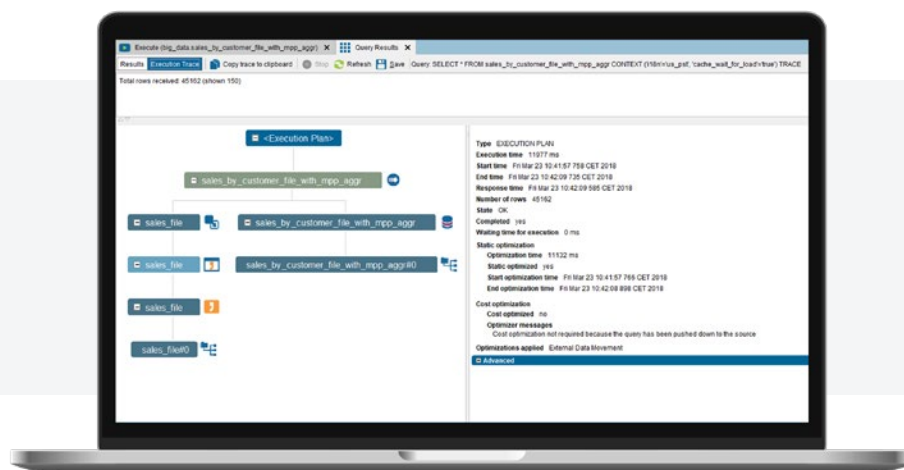
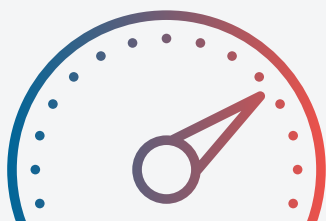
Denodo 6.0は業界最高の動的クエリ最適化エンジンを提供しており、これには、全集約・部分集約プッシュダウン、オンザフライのデータ再配置、代替データソースの使用など、ロジカルアナリティカルアーキテクチャのネットワークトラフィックを最小化する独自の技術が採用されています。Denodoの高度なコストベースオプティマイザは、中間結果セットのサイズ、データソースインデックス、データソースの実行モデル(パラレルデータベース、Hadoopクラスタなど)、データソースの特性(パラレルデータベースまたはHadoopクラスタの処理コア数)、ネットワーク転送速度を統計情報に基づいて推定する業界唯一の最適化エンジンです。

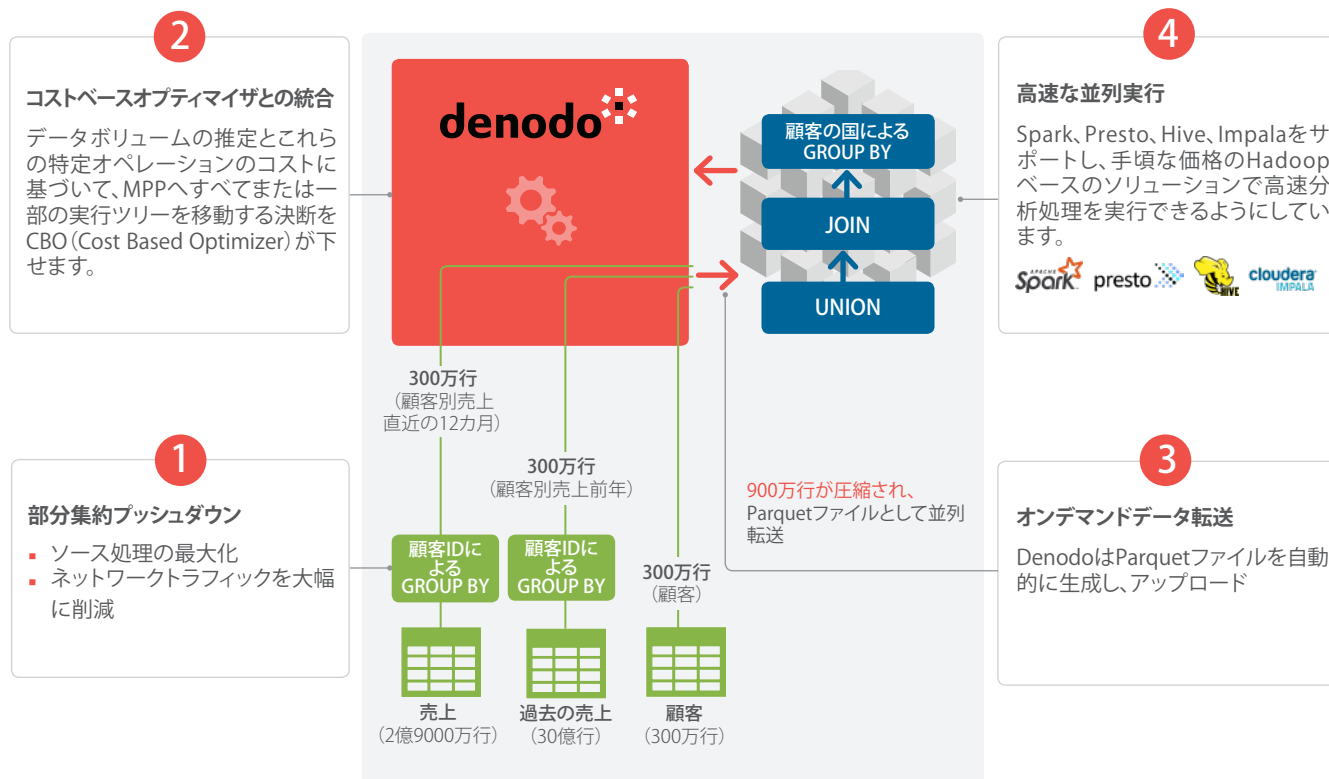
Denodo Platform 7.0は、複雑なビッグデータの環境で最高のパフォーマンスを発揮し、データ仮想化ソリューションの限界を押し上げます。これには、クエリを加速化するMPPエンジンのネイティブ統合も含まれます。Denodo Platform 7.0は、一般的な並列処理技術(Spark、Hive、Impala、Prestoなど)に基づいているMPPクラスタのDenodo内の操作を並列処理し、作業負荷が非常に高い環境でもクエリの実行を加速化できます。これに加えて、Denodo Platform 7.0はMPPキャッシングを提供し、これにより最近使用したデータへ高速にアクセスすることができます(Spark、Hive、Impala、Prestoが現在キャッシュシステムとしてサポートされています)。

動的クエリ最適化とMPPクエリ加速化の両方の技術を組み合わせて、ビッグデータボリュームのある非常に複雑な環境でも、手頃な価格のMPPソリューション(Spark、Presto、Impalaなど)を利用して、非常に効率の良い方法で問題なくアプローチすることができます。

Denodo 7のクエリ最適化は複数の段階から構成されています。まず、高い処理能力を利用し、データソースからのデータ移動を最小化することにより、クエリをデータソースに最大限委譲できるようにします。次に、Denodoで結合される必要のあるデータソースから受信された部分的データセットについて、(Denodoのコストベース最適化モデルに基づいて)MPPクエリの加速化により処理速度の向上を図ります。

Denodoの動的クエリオプティマイザは、データソース(Impala、Spark、Teradata、Netezza、Verticaなど)の大規模な並行実行能力を委譲または活用するための最適な戦略を自動的に選択します。クエリ処理をデータソースにプッシュダウンする場合に最適なパフォーマンスを実現するために、処理を分散させるインテリジェント「マスター」の役割を果たします。部分的な結果がデータソースで計算されると、Denodo Platformは、元のデータセットよりはるかに小さいデータセットで要求される最終的なデータ結合および変換を行います。MPPクエリの加速化メカニズムにより、処理をSpark、Impala、またはPrestoのいずれかに移動させて並列に実行し、最終処理の加速化を図ることができます。処理の加速化を実現するために、システムはParquetファイルを自動生成し、データをMPPクラスタにオンザフライで転送します。これは、メモリ内の最終処理を実行するか、または処理をMPPクラスタに委譲するかを最終決定するために使用されるDenodoのコストベース最適化モデルを基盤としています。データ仮想化エンジンは、いつデータがMPPクラスタ内でローカルとなっているか、または不要なデータ移動を回避するためにあらかじめキャッシュされているかを自動的に検出します。





システム	実行時間	最適化技術
Denodoなし	～10分	シンプルなフェデレーション
MPPなし	43秒	集約プッシュダウン
MPPあり	11秒	集約プッシュダウン + MPP統合 (Impala 8 ノード)

表でハイライトされている実行時間からもわかるように、Denodo 7.0はパフォーマンスを顕著に向上させました。上の図では、大規模なデータの集合体に対して集約オペレーションが行われています。Denodoが、最初の部分集約プッシュダウンを実行し、ネットワークトラフィックを約6億行から900万行に減らしています。それがネットワークを介してデータ仮想化レイヤーに転送されます。Denodoはその900万行の処理をMPPクラスタに移行します。そこでは、複数の圧縮形式のParquetファイルでデータが並列転送されます。実行時間は、基本的なクエリ最適化技術を適用した場合は10分を超え、Denodo部分集約プッシュダウンを適用した場合は43秒まで下がり、そして部分集約プッシュダウンとMPPクエリ加速を適用した場合は11秒まで下がります。データの一部がMPPエンジンにキャッシュされているか、MPPに既にあるローカルテーブルからのものである場合、最適化はこの状況を考慮して、不必要なデータの移動を回避します。

つまり、Denodo 7.0は高度なクエリプッシュダウンを応用して、大量の並行処理を行うネットワークトラフィックを最小化します。

さらに、新しい動的クエリ最適化が統合されていて、部分集約プッシュダウンのためにプライマリキーが必要なく、Amazon S3データソースのためにソースからソースへ直接データ移動を行う必要もありません。

## 動的データカタログ

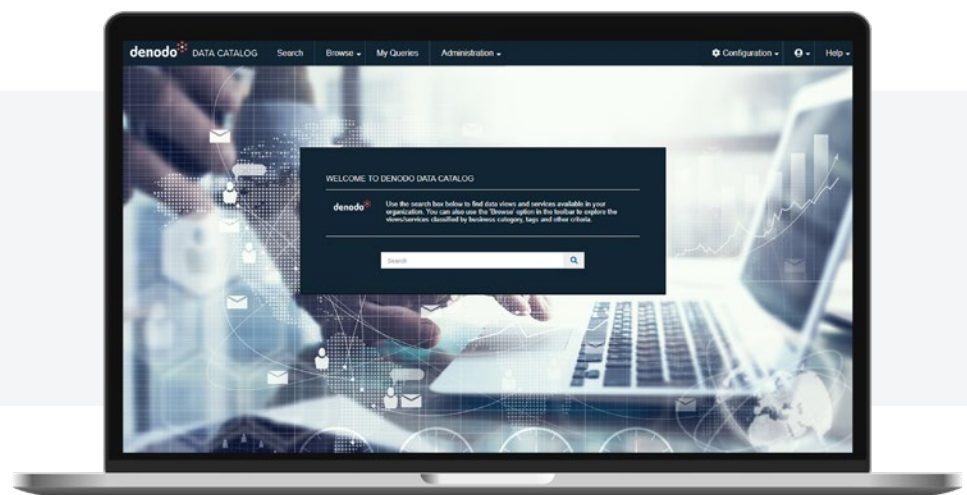
セルフサービス戦略の主な特徴として、ビジネスユーザ（シチズンアナリストと基幹業務アプリケーション開発者の両方）が、データ転送レイヤーでどのデータセットを使用できるかを確認し、自分達が必要としている情報にどれが関連しているかを確認できるようにすることを挙げるすることができます。

今回新しくリリースされるDenodo Platform 7.0には、DenodoデータカタログというDenodo 6.0インフォメーションセルフサービスツールを進化させ、ブランディングを刷新した機能があります。これは、ビジネスユーザがセルフサービスでデータの検出やブラウジングの処理を行えるように支援することを目的としたツールです。Denodoデータカタログがあるため、ビジネスユーザと基幹業務アプリケーション開発者は、Denodoデータ仮想化レイヤーを通じて利用できる全データセットの検索、共有、再利用を行えます。ビジネスのデータスチュワードは、Denodoのメタデータに基づいてビジネスビューのカタログを作成し、ビジネスカテゴリに従ってそれらを分類し、それらをタグ付けしてアクセスしやすいようにできます。ユーザは、ビジネスのカテゴリやタグを通じてブラウジングやナビゲーションを行い、データセットの関係を発見し、メタデータやデータそのものを検索することにより、データセットを適切に検証し、ツリービューや詳細なリネージ情報を点検し、クエリの結果をエクスポートします。

Denodoが提供するデータ転送インフラストラクチャと密接に連携しているため、メタデータが受動的にしか更新されないためにすぐ使われなくなってしまう従来の静的なカタログと異なり、能動的に最新の状態に保たれ続ける動的なカタログとして保持されます。Denodoは、データカタログをデータ転送とシームレスに統合する業界で唯一のデータ仮想化ソリューションです。Denodoデータカタログは、キュレートされているタイムリーでコンテキスト依存の再利用可能な情報資産とデータサービスの動的カタログです。

ユーザは、Denodo Platformの強力な検索機能とブラウザ機能を活用し、メタデータだけでなく実際のデータにアクセスでき、これにより利用可能なデータセットの理解を深めることができます。さらに、ユーザは、利用可能なデータセットに対して自分独自のカスタマイズを行うことができます。ユーザは自分独自のデータ変換をグラフィカルウィザードで構築できます。ウィザードには、JOINや集約などに関する複雑なクエリが含まれ、Denodo Platformにより内部で自動的に実行されます。ユーザは複雑なSQLコンストラクトを知る必要はありません。複雑なことは隠蔽され、ビジネスユーザが使いやすいWebベースのインターフェースで操作を行えます。

ユーザは後日クエリを使用できるように保存し、リンクを共有するだけでユーザとそれらを共有できます。また、その他のビジネスユーザ向けに展開する前に、IT部門と共有するサンドボックス環境へエクスポートすることで事前検証が行えます。これによりビジネスユーザとIT部門のシームレスなコラボレーションを確保できます。

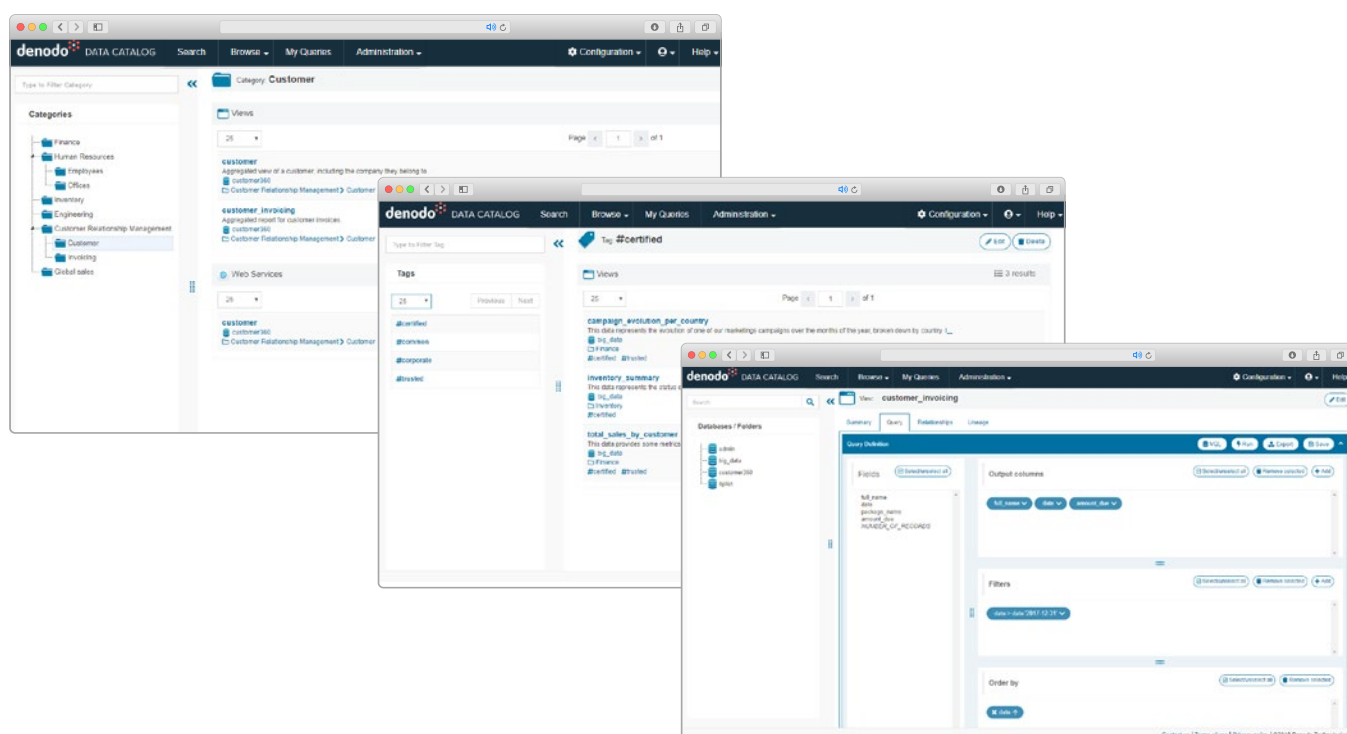


データ配信レイヤーをデータカタログと組み合わせることにより、コンテキスト化されたデータ資産利用を行うことができます。誰が何をいつどのデータセットを使っているかといった、関連性と時系列を示すインフォノミクスの基準（Denodoデータ転送インフラストラクチャとの自動統合により、どの本番アプリケーションが現在一定のデータセットを使用しているかなど）を生成することを支援できます。

Denodoデータカタログを利用すると、デジタルマーケットプレイスによりアナリストや意思決定者のコミュニティが情報資産を迅速に検索し、使用できるようになります。これはセルフサービスアナリティクスが求められる現在の環境で不可欠なことです。

Denodoデータカタログは外部データガバナンスやカタログツールを補完するものであるため、これらと併用できます。Denodoデータカタログはビジネスユーザやアプリケーションによるデータ活用のために公開されるデータセットのカタログです。一方、データガバナンスとカタログのツールはITインフラストラクチャ全体のカタログを構築することにフォーカスしています。

Denodo Platform 7.0も、インストールしてすぐに利用できる外部データガバナンスツールとの統合機能をDenodoガバナンスブリッジを通じて確保できます（IBM InfoSphere Information Governance Catalogがリリース時にサポートされる最初のツールです）。



## ITペルソナとビジネスペルソナ

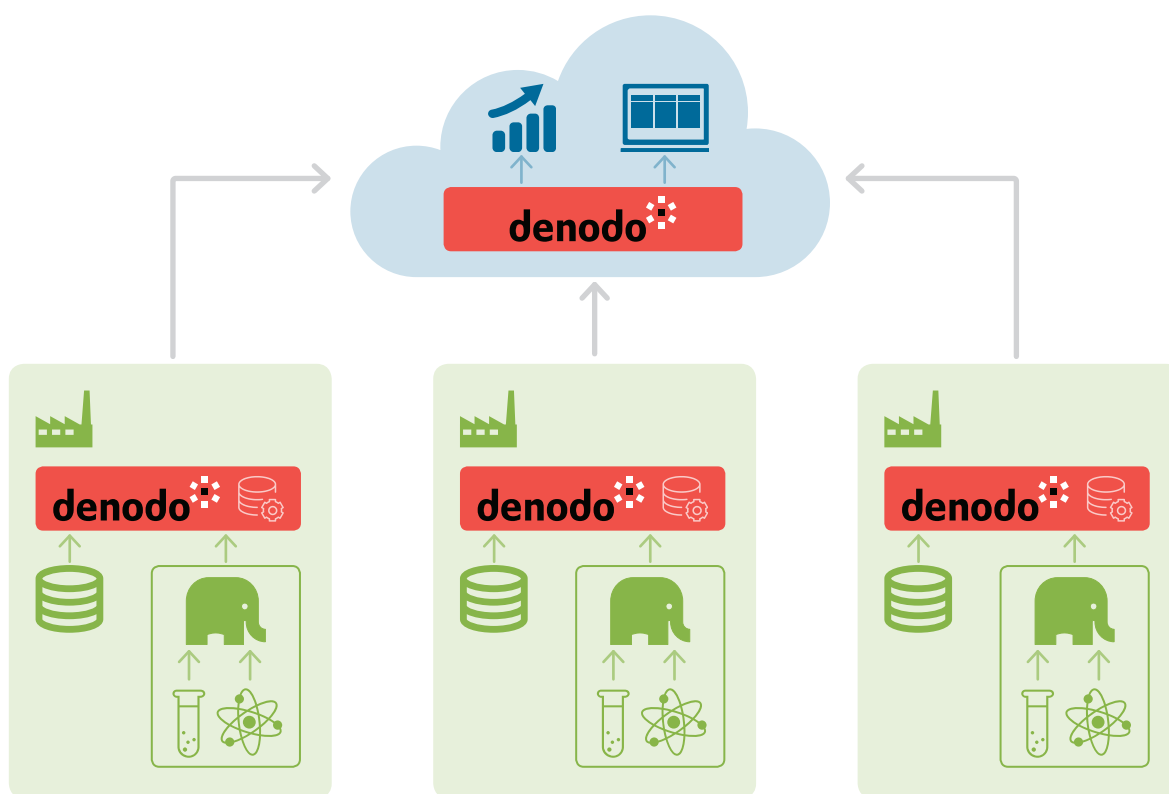
Denodo Platform 7.0は、データ指向のユーザとビジネス指向のユーザ両方に対してインターフェースを提供します。

- データエンジニア、パワーユーザ、シチズンインテグレイタなどのデータ指向の開発者に対して調整されているドラッグアンドドロップやコーディングレスの開発機能のある統合開発環境で、マウスを数回クリックするだけでデータサービスを公開できます。
- 改善されたデータカタログのUI - データスチュワード、データアナリスト、シチズンデータアナリストなど、ビジネス指向のユーザに対して簡単に使用できるインターフェースを提供します。
- ホリスティックソリューションマネージャのUI - デプロイメントやプロモーションを管理し、タスクを調整したり、モニタリング、監査、トラブルシューティングを行ったりする際、管理者に一元化されたUIを提供します。

## マルチクラウド、ハイブリッド、エッジによる、柔軟性の最大化

Denodo Platform 7.0は、マルチクラウド、ハイブリッドおよびエッジのシナリオで、パフォーマンス、セキュリティ、ガバナンスにおいて妥協することなく、マルチロケーションアーキテクチャを提供します。このソリューションは特にロケーションの透過性を確保するために設計されており、コストのかかるロケーション間データ移動を最小限に抑えます。Denodoマルチレイヤーアーキテクチャを利用すると、データソースの近くの同じ場所に配置できるDenodoインスタンスの階層が、リモートロケーション（オンプレミス/クラウド）でローカル処理を実行できるため、ネットワーク上のデータ転送を激減できます。ローカルキャッシュは、ローカルノードにおいて後続のクエリで利用できる直前に計算されたデータを保存できます。Denodoマルチレイヤーアーキテクチャの別の優位性として、ファイアウォールの制限をバイパスするのを支援するため、セキュリティを向上できることを挙げるすることができます。

Denodoソリューションマネージャは、マルチロケーションアーキテクチャをサポートする上で重要な役割も担っています。Denodoのリモートインスタンスからのメタデータを自動的に適合し、階層の各ノードの関連データソースを指し示すようにし、階層の高レベルのDenodoインスタンスメタデータを適合し、低ノードの適切なDenodoインスタンスを指し示すようにします。



Denodo Platform 7.0は、AWS、Azure Marketplace、Dockerコンテナの業界唯一のデータ仮想化プラットフォームで、ユーザが小さく始めて後日拡張できる動的かつ柔軟なインフラストラクチャを提供します。この目的のためにDenodo Platform 7.0は、AWSとAzureの自動拡張機能をサポートし、作業負荷に従って新しいサーバをクラスターに動的に追加できるようにしています。

Denodo Platform for AWSとAzureは、Denodo Platform完全版と同じです。このソリューションは、エンタープライズ級のセキュリティ、最先端のパフォーマンス最適化、データガバナンスと管理、ならびにデプロイメントモニタリング、管理、そしてセルフサービスで対応できるデータリカバリ機能を提供します。これは、設定が必要なく、柔軟な時間単位のリースオプション、幅広い容量オプションを提供し、あらゆるビジネスニーズに対応します。

Denodo Platform 7.0は、Dockerによるコンテナリゼーションをサポートしており、アプリケーションデリバリーを加速化できます。



# 自動ライフサイクル管理

Denodo Platform 7.0は継続的な転送を行うための新しいツールとAPIで開発運用を合理化します。新しいDenodoソリューションマネージャは、複雑なマルチクラスタデプロイメントを一元化されたWeb UIを通じて行うことができます。これにより、シンプルで効率の良い管理とデプロイメントを行うことができます。これは、アジャイル方法論と継続的デリバリーポリシーのメリットを実現する上で非常に重要です。

Denodoソリューションマネージャのツールを利用すると、次のような複雑なタスクを自動化できます。

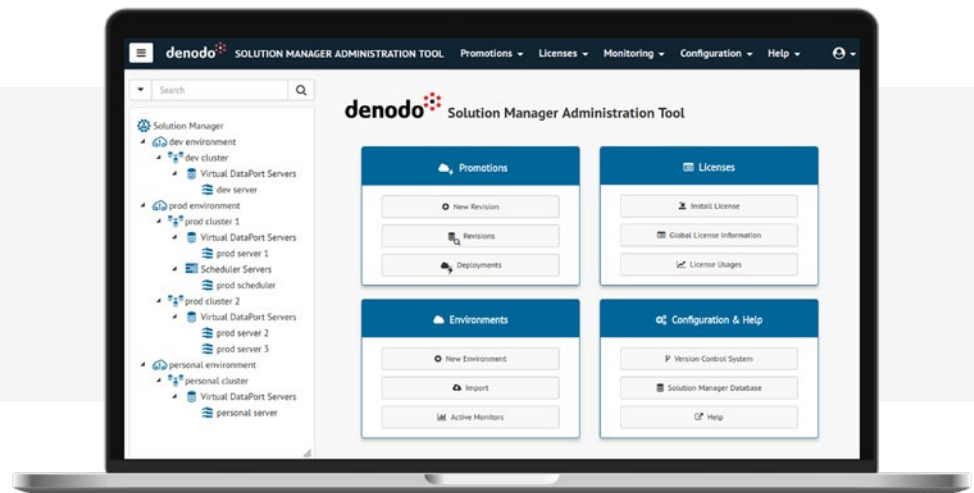
- ・ ライセンスキーの一元管理
- ・ データソースの環境依存プロパティ(ログイン、パスワード、ホスト名など)の一元管理
- ・ Denodoクライアントへの更新の管理と配信の一元化
- ・ 別のクラスタ(オンプレミス/クラウド)のDenodoインスタンス間における新しいメタデータ更新の自動プロモーション
- ・ ダウンタイムがゼロの環境の間の新しいメタデータの自動プロモーション(開発から品質保証、あるいは品質保証から本番など)
- ・ 利用できる全インスタンス間におけるログの自動管理
- ・ DenodoソリューションマネージャはWebベースツールで提供されるため、すべての保守サポートはグラフィカルインターフェースを利用して行われます。

グラフィカルインターフェースを利用する保守サポートに加えて、新しいリリースにはREST APIがあり、これによりJenkins(例えば、Denodoで定義されたマイグレーションはJenkinsで管理できます)、uDeploy、Chef、Apache Builderなどの開発運用ツールからのタスクを自動化できます。

また、ソリューションマネージャは、各環境(ステージング、本番)でアクティブだった各バージョンのコピーを保存し、後日分析したり、バックアップを確保したりできるようにします。使用量とパフォーマンスに関するデータを提供するための情報モニタリングがあり、ガバナンスを支援することもできます。

さらに、ソリューションマネージャを利用すると、検証を自動化できる(例えば、データソースのプロパティがまだ追加されていないときに、データソースを本番にマイグレーションしようとするユーザーが警告を受けます)ため、提示されたビューが特定の制限を検証しないと、ビジネスユーザーは警告を受け取ることができます。また、このソリューションを使うと、IT部門が、提示された要素を確認し、それをステージング環境でテストして変更が問題ないか確認し、最後に本番環境に公開することができます。

Denodo Platform 7.0を利用すると、デプロイメント、更新、ロギングなどの自動ライフサイクル管理機能の全ての開発運用タスクを劇的に簡素化でき、ビジネスに必要な迅速性を実現できます。



# データサービス公開、ガバナンスとセキュリティー、ならびに 接続性、データ変換の新機能

## データサービス公開

クラウドの相互運用性を簡単に確保するために、Denodo Platform 7.0はモダンデータサービスレイヤーを提供し、最新の規格を実装しています。JSON Web TokenやOAuthトークンのイントロスペクションの公開されたデータサービスのOAuthサポート、公開されたWebサービスに対するドキュメント自動生成に対応するためのオープンAPI仕様のサポート、OData 4インターフェースの機能拡張などです。

## ガバナンスとセキュリティー

Denodo 7.0は、データソースとデータサービスに対する許可や、新しい「メタデータ」権限など、カタログ内の全要素に対してきめ細かい権限を追加しています。これによりビューのスキーマ、ツリービュー、実行トレースが確認できます。

Denodo Platform 7.0は、ユーザが割り当てた使用クォータを設定することを通じて、IT部門がデータサービスレイヤーを制御する新しい方法をさらに提供します。使用クォータは、Denodoリソースマネージャにより(分/時間/日/月単位で)定義できます。

新しいレポートが追加され、ユーザアクティビティの追跡機能が改善されました。また、新しい「pingデータソース」機能により、全データソースの現状をチェックできるようになりました(モニタリングと診断ツールのグラフィカルインターフェースとAPIの両方で利用でき、ユーザはそれに基づいてアラートを生成できるようになります)。

## 接続性

Denodo Platform 7.0には、SalesforceとPresto DBの新しいアダプタがあります。新しいSalesforceのアダプタは、パフォーマンスを最適化するために、改善されたグラフィカルな管理機能と、改善されたプッシュダウン機能を提供します。Presto DBは、データソース、キャッシュシステム、MPP加速化エンジンとしてサポートされています。以前のバージョンと比べると、Presto DBに対するクエリプッシュダウン機能が大幅に改善されています。

DenodoはTableauやMicrosoftと協力してTableauやPower BIのアダプタを開発し、相互運用性を保証できるようにしています。

## データ変換

最後に、Denodo Platform 7.0はサポート対象のデータタイプをSQL標準の新しい日付/時間タイプで拡張し、地理空間データの機能を新しい地理空間機能で拡張しました。異なる座標参照系(CRS)のサポートと、GeoJSON規格に準拠したデータサービスの公開です。



Denodo Technologiesは、アジャイルな高パフォーマンスデータ統合、データ抽出、リアルタイムデータサービスを、企業、クラウド、ビッグデータ、非構造化データソースの範囲のすみずみにわたって、従来のアプローチの半分のコストで提供するデータ仮想化ソリューションのリーダーです。Denodoの顧客は様々な主要産業にわたり、これらの顧客はビジネスできわめて迅速な対応ができるようになり、非常に高いROIを実現しています。