

Spotlight by Bloor

Andy Hayler

2024年1月

データファブリック 概要



データファブリックは、ソースデータのコピーを移動させることなく、企業内データの統一ビューを得るためのアプローチです。代わりに、ソースデータをマッピングする中核となるセマンティックレイヤーが作成され、それをナビゲートしてクエリできるようになります。

はじめに

データファブリックとは、データ管理のためのアーキテクチャのひとつで、データを記録し、ユーザーからのクエリに応じてオンデマンドでデータを提供するソフトウェアの仮想レイヤーを意味します。データウェアハウスなどのアプローチとは異なり、データを中央のハブに集めるのではなく、その場に残しておく代わりに、クエリはオンデマンドで分散して実行されます。

データファブリックは成長分野であり、データカタログからナレッジグラフ、人工知能に至るまで、数十のベンダーがこれをサポートすると主張しています。現時点では、業界はアーリーアダプターの段階にあるようです。

ベンダーによる盛り上がりとマーケティング活動がありますが、顧客の声、定量化されたビジネス上の利点を持つ匿名ではない企業の、文書化されたケーススタディは少ないといえます。Bloorがデータファブリック業界の現状を調査し、ソフトウェアのパワーポイント層の下にある現実に迫ります。

“データファブリックは、データの統一ビューを得るためのアプローチであり、マスターデータ管理、データ品質、データメッシュ、データウェアハウジングとは異なりますが、これらの分野とも関係があります。”

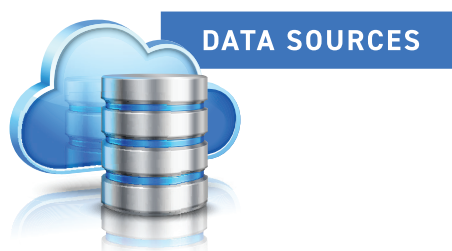
データファブリックとは

グローバル企業は長い間、データ資産の管理に苦労してきました。一般的な企業は、オンプレミスのデータセンター、AzureやAWSのようなパブリッククラウド、プライベートクラウド、それ以外にも散在する数百のアプリケーション（調査によって数字は異なりますが、いずれも数百に上る）を抱えています。長年にわたり、データの群れをまとめるために、多くの試みがなされました。企業は、データウェアハウス（分析に必要なものを中央の場所にコピーする）、マスターデータ管理（顧客、製品、場所などのアプリを横断する主要データを管理）、エンタープライズリソース・計画（すべてのアプリを単一ベンダーの巨大アプリに置き換える）を試してきました。こうした取り組みに巨額の投資が行われたにもかかわらず、エンタープライズアプリケーションの数は着実に増加しており、完全かつ普遍的に成功したと証明されたものはないと言っていいでしょう。複数の独立機関による調査によると、2023年時点で自社のデータを信頼している経営幹部はわずか3分の1であり、これは数年前とほぼ同じレベルです。

データファブリックとは、企業のデータリソース全体を整理・監督し、すべてのシステムとアプリケーションが必要とする、すべてのデータとサービスにタイムリーにアクセスできるようにするデータアーキテクチャとインフラストラクチャのことです。構造化データと非構造化データの両方にまたがり、既存のメタデータを利用することもあります。ちなみに、これは「データメッシュ」、分散化されたデータ所有権によって主要なデータ領域に焦点を当てた、データ管理に対するほとんど組織的なアプローチ、とは異なります。



BUSINESS USERS



“データファブリックとは、企業のデータリソース全体を整理・監督し、すべてのシステムとアプリケーションが必要とする、すべてのデータとサービスにタイムリーにアクセスできるようにするデータアーキテクチャとインフラストラクチャのことです。”

なぜ注目すべきなのか

データウェアハウスやデータレイクのような新しいデータソースに、大量のデータを事前に統合したりコピーしたりすることなく、企業全体のデータを可視化し、照会する方法だからです。中央データカタログは、企業全体のデータの宝の地図のようなもので、ユーザーはデータの物理的な場所を知らなくても「顧客」「製品」「資産」などのビジネス用語を使用してデータをナビゲートできます。データファブリック・テクノロジーは、データの矛盾や重複をリアルタイムで解決できるものでなければなりません（ある調査によると、平均的な企業では「顧客」の「マスター」定義が6つ、「製品」の定義が9つあるとのこと）。

データファブリックは、必ずしも既存のデータ管理構造をすべて置き換えるものではないことに注目すべきです。例えば、マスターデータ管理ハブは、データファブリック・アーキテクチャにおける別のソースシステムに過ぎません。データファブリックの仮想レイヤーが、ほとんどの企業における、様々な顧客階層や製品階層間の不整合を、どのように動的に解決するのかは不明なため、マスターデータ管理ハブは非常に役に立つかもしれません。MDMハブは、様々なソースの信頼性を分類する「生存ルール」を持ち、データの不整合を解決するために、データ品質ルールと統合/マッチングアルゴリズムを適用します。既存のMDMハブにプラグインするデータファブリック仮想レイヤーによって、膨大な作業をすべて利用することができます。そうしたハブが存在しない場合、データファブリック・テクノロジーは、このような問題に自ら対処しなければなりません。

“ データウェアハウスやデータレイクのような新しいデータソースに、大量のデータを事前に統合したりコピーしたりすることなく、企業全体のデータを可視化し、照会する方法です ”

メリット

データファブリック・ソリューションの導入に成功した場合、多くのメリットを享受できる可能性があります。例えば営利企業であれば、様々な企業とのタッチポイントにおいて、顧客に関する統一されたビューを得ることができます。病院は、患者の記録や治療履歴などを含む、患者に関する単一で一貫したビューを得ることができます。金融機関は、各取引における真の取引相手が誰であるかを効率的に把握することで、取引リスクの評価を向上させることができます。

例えば、最も収益性の高い製品や顧客が誰であるか等、単純ですが、実際に情報提供するのは難しいことです。基本的に包括的なデータファブリックは、ビジネスとともに成長し、新しいデータソースやビジネスモデルの変更に適応します。もしそれが予定通りに機能するならば、頑固で根絶が難しい、もうひとつの茨の道であるデータ品質も改善されるはずです。

このようなアーキテクチャを実際に機能させるには、特に企業規模では大きな課題があります。大規模な単一データベースに対して複雑なクエリーをリアルタイムで実行するには、かなりのリソースを必要とするからです。データが分散した複数のデータベースに対して、そうしたクエリーを実行するのは困難な問題であり、非常に効率的な分散クエリー機能を必要とします。現在主流となっているデータベースで、このような能力があると本気で主張できるものはほとんどないでしょうが、専門的なニッチデータベースは分散クエリーを念頭に設計されているものもあります。

たとえその問題が無視されたとしても、データ品質の問題があります。これは、さまざまなデータ品質ソフトウェアやデータガバナンス・プロセスで何十年にもわたって対策が試みられてきたにもかかわらず、運用システムを悩ませ続けている問題です。経営幹部のデータに対する信頼度は、2023年の調査でも依然として低いままで、クエリーを実行できるだけでも意味があります。データウェアハウスやマスタータデータ管理といった従来の、品質にかなりの重点が置かれていました（典型的なマスタータデータ管理プロジェクトの約3分の1は、データ品質問題の解決に費やされています）。データファブリックは、そのようなハブへのデータの移動を拒否し、クエリーを動的に解決することを好みますが、根本的なデータ品質に関する非常に現実的な問題を、どのように取り除くことができるかは不明です。確かに、どのようなデータファブリック・イニシアチブであっても、データ品質の課題にどのようにうまく対処するのかを説明する必要があります。私がそうであったように、データファブリックのケーススタディを探すのに時間をかければ、ベンダーの主張はたくさん見つかるでしょうが、顧客の声や定量化されたビジネス上の利点を持つエンドユーザー企業の名前を挙げているものは、ほとんどありません。データファブリックは新しいアイデアではありませんが、適切に文書化されたケーススタディが少ないことから、まだアーリーアダプターの段階にあると思われます。データファブリックのベンダーやコンサルタントが宣伝している匿名の事例も多く、彼らは明らかに既得権益を主張しています。



経営幹部のデータに対する信頼度は、2023年の調査でも依然として低いままで、クエリーを実行できるだけでも意味があります。



ソリューション

データファブリックにはいくつかの基本コンポーネントがあります。ソース・アプリケーション全体のデータ資産をマッピングする仮想ソフトウェアレイヤーが確立され、ソースデータはそのままに、ビジネスニーズを満たすために必要に応じてデータを変換・処理しようとします。その中核となるのは、ビジネス用語集を含むデータカタログ、何らかの形のレコメンデーションエンジン（AIコンポーネントを含む）、データランドスケープを視覚化する方法（データエンティティのリンクされた説明の視覚的集合である「ナレッジグラフ」など）、必要に応じてデータを取得するデータ準備レイヤー、オンプレミスやクラウドを含むデータのエンドツーエンドのワークフローを管理するデータオーケストレーション・コンポーネントです。

現在、何十社ものベンダーがデータファブリック・ソリューションを持っていると主張しています。その中には、データ管理ソリューションの長い歴史を持つ巨大ベンダーもあれば、データカタログから人工知能「アクティブ・メタデータ」まで、さまざまな機能を持つ新興企業もある。全コンポーネントをカバーするエンドツーエンドのソリューションを持っているところは少ないですが、そう主張しているところもあります。

“ 今日、何十ものベンダーがデータファブリック・ソリューションを持っていると主張していますが、すべてのコンポーネントをカバーする、エンドツーエンドのソリューションを持っているところはほとんどありません。 ”

まとめ

データファブリックは、ソフトウェアベンダーが、データサイロを増やすことなく運用中のソースシステム内のデータに動的にアクセスすることができるという、柔軟なデータアーキテクチャの利点を、顧客が実現できるよう支援するためにマーケティング活動を展開している新しいアーキテクチャです。しかし、このようなアプローチが直面する現実的な課題も少なくありません。データウェアハウス、データレイク、マスターデータ管理といった以前のアプローチが苦戦を強いられた根本的な問題は消えておらず、データファブリックが、こうした以前のアプローチとは真に異なる布地から切り出されるためには、根本的な問題を解決する必要があります。

著者について

ANDY HAYLER

Practice Leader: Data as an Asset



ソフトウェア業界の権威であり、独立系戦略コンサルタントとして企業、ベンチャーキャピタル、ソフトウェア会社にアドバイスを行う。カリドの創業者。彼のリーダーシップの下、カリドは2001年に世界で最も急成長したビジネスインテリジェンス・ベンダーとなった。カリドはデータウェアハウスのイノベーターとして認められ、当時はまだ存在しなかったが、現在ではよく知られ急成長している業界で、最初の真のマスターデータ管理製品を発売した。

アンディは、レッドヘリングの「2002年イノベータトップ10」にヨーロッパ人で唯一選ばれた。受賞歴のある「Andy On Enterprise Software」ブログは、ブログのパイオニアでもある。

アンディはエッソでキャリアをスタートさせ、多くの技術職を経験した後、シェルに移った。シェルUKのテクノロジー・プランニング・マネージャーを経て、シェル・インターナショナルのプリンシパル・テクノロジー・コンサルタントを務めた。その後、グローバルな情報管理コンサルタント会社を設立し、彼のリーダーシップの下、300人のスタッフを抱えるまでに成長した。

Bloor概要

テクノロジーがビジネスの急速な進化を可能にしています。チャンスは計り知れないものの、適応しなければ生き残ることはできません。変化しやすいビジネスの時代には、進化が成功に不可欠なのです。

私たちは未来を示し、実現するお手伝いをします。

Bloorは、お客様が複雑なビジネス状況に対応し、課題を真の成長、収益性、影響への新たな機会に変えることができるよう、斬新な技術的思考を提供します。

革新的な独立系テクノロジーリサーチ、アドバイザー、コンサルティングサービスを通じて、実用的な戦略的洞察力を提供します。私たちは、複雑なビジネス状況に斬

新な思考をもたらし、課題を真の成長と収益性のための新たな機会に変えることで、企業の変革の旅を通して、常に適切であり続けるよう支援します。

Bloorは25年以上にわたり、戦略を調整し、可能な限り最高の結果を達成するためにテクノロジーを取り入れることによって企業のインテリジェントな進化を支援してきました一貫して改善し成功するために、前提に挑戦するお手伝いをします。

著作権および免責事項

書の著作権は Bloor 2024 に帰属します。本書のいかなる部分も、プロア・リサーチの事前の承諾なしに、いかなる方法によっても複製することを禁じます。

本資料の性質上、多数のハードウェアおよびソフトウェア製品の名称が記載されています。すべてではないにせよ、大半の場合、これらの製品名は製品を製造している会社によって商標として登録されています。これらの名称または商標を当社のものとして主張することは、プロア・リサーチの意図するところではありません。同様に、会社のロゴ、グラフィックまたはスクリーンショットは、所有者の同意を得て複製されたものであり、その所有者の著作権の対象となります。

本書の作成にあたり、情報が正確であるよう細心の注意を払っておりますが、誤記や脱落があった場合、発行者は責任を負いかねます。



Bloor Research International Ltd

-  20-22 Wenlock Road, London N1 7GU, United Kingdom
-  +44 (0)1494 291 992
-  info@Bloorresearch.com
-  www.Bloorresearch.com