



# サードパーティデータをアナリティクスで利用するための ベストプラクティス

管理された方法で外部データへライブアクセスすることにより、データエンリッチメントと新しいインサイトの獲得をスピードアップ



# TABLE OF CONTENTS

- 2** エグゼクティブサマリー
- 3** 外部データソースはデータアナリティクスに不可欠
- 4** 秘められたデータの価値を解き放つ3つの方法
- 6** サードパーティデータの調達は容易でない
- 7** データクラウドで従来のデータシェアリングの課題を解決
- 8** Snowflakeデータマーケットプレイス:安全なデータシェアリングの活用
- 9** 今すぐデータコンシューマーとしての能力を高めましょう
- 10** Snowflakeについて

# エグゼクティブ サマリー

社外のソースから取得するサードパーティデータの活用により、企業はデータアナリティクスを改善し、インサイト取得までの時間を短縮することができます。外部データはチームがよりデータドリブンな意思決定を行うための有用なリソースであり、また内部データとの統合によりその価値は一層高まります。しかし、従来のサードパーティデータの調達方法は非効率的で、セキュリティが不十分です。従来のデータマーケットプレイスは拡張性が不十分で、古いデータ転送技術（FTPやAPI）では大規模なエンジニアリング作業を行わなければデータを利用できません。そのため、遅延、データの鮮度低下、劣悪なデータアナリティクスといった問題が生じます。

本eBookでは、以下の方法について説明します。

- 元の保管場所から移動させることなくサードパーティデータにライブアクセスすることにより、即時に分析への利用や自社データとの統合が可能となる
- 自社のビジネスニーズにあったサードパーティデータを簡単に見つける
- 自社データをプロバイダーと安全に共有してエンリッチメントサービスを利用することにより、内部データ品質を改善する



# 外部データソースは データアナリティクスに不可欠

目まぐるしく移り変わる昨今のビジネス環境において、企業の強さは利用するデータの有効性に比例します。10年以上の間、多くの企業は自社データの収集、保管、分析に多くの時間、資金、リソースを投じてきました。

内部データの分析により経営や顧客との関係を改善する方法が明らかになる場合がある一方で、ほとんどの企業は市場の変化の把握、競争力のある新しいインテリジェンスや既存のイ

ンテリジェンスの獲得、消費性向の監視に必要な幅広いデータを取得できていません。

足りないのは、外部データです。サードパーティデータは、企業が社内で生成するデータの限界を超えて、データに基づいたインサイトを獲得するために役立ちます。そのような幅広いデータを分析することにより、企業は広範な業界トレンドへの理解を深め、競合他社が対応するより早く市場の変化を把握し、新製品や新サービスの提供に向けて顧客の行動を分析し、正確に売上を予測することができるでしょう。

戦略や財務に関するあらゆる意思決定がデータ分析にかかっているこの時代に、**より多くのデータを利用しない理由はありません。**

サードパーティデータで内部データを強化することの重要性が、かつてないほど高まっています。Forrester社のレポートによると、世界中のデータおよびアナリティクス担当の意思決定者

の56%は外部データ調達能力を強化済み(または強化中)であり、21%は今後12か月間でそれを実行する予定です。<sup>1</sup>また、Gartner社は、正式なデータマーケットプレイスでデータを売買する大企業の割合が2022年までに全体の3分の1以上(35%)に達し、2020年と比較して25%増加すると予測しています。<sup>2</sup>

個人、企業、センサーなどのさまざまなソースによって生成される外部データには、以下のような幅広いカテゴリーがあります。

- オンライン行動(検索、ソーシャルメディア、アプリ利用、ウェブトラフィック、地理位置情報)
- 消費者(取引、消費者マインド)
- 個人(就労、信用)
- 企業(広告、価格、評価およびレビュー、店舗所在地)
- イベント(売買、衛星および気象、イベント検知)
- 集合データ(IoT、ウェブクローラ、B2B、オープンデータ)

新たなソースから関連データを取得することで、企業はデータアナリティクスの改善、データモデリングの強化、インサイト獲得までの時間の短縮といったメリットを得ることができます。また、サードパーティデータは、外部検証を行わなければ証明するのが難しい仮説を検証するために役立ちます。

<sup>1</sup> bit.ly/2Fh8WzF

<sup>2</sup> gtnr.it/3nBrc8j

# 秘められたデータの価値を 解き放つ3つの方法

データベンダーから取得した外部データを利用することにメリットがあるのは明らかですが、企業は外部データを内部データやアナリティクスと統合することでさらなる価値を引き出すことができます。

以下では、業界を問わず、消費者の行動、財務状況、世界的な事件の影響を把握する上で外部データがどのように役立つのかを示す3つの例を紹介します。

## 内部データのエンリッチメントにより、ビジネス上の意思決定を改善

消費者向けの製品やサービスを提供している企業は、多くの場合、内部データを利用して店舗のトレンドを把握します。たとえば、小売業者は店舗の販売記録を分析し、店舗の実績と販売予測を比較します。それでは、気象パターンが店舗の売上に影響しているかどうかを確認するために小売業者が過去の気象データと内部データを組み合わせて利用したらどうなるかを考えてみましょう。過去のトレンドを調べるだけでなく、リアルタイムの気象予測データを利用して、もうすぐやってくる雷雨や猛暑の影響を受ける可能性があるのはどの店舗かを予測することができます。この分析は、悪天候に見舞われた地域の販売予測を修正するための指標となります。

また、人の移動データも、有用なサードパーティデータです。店舗の近くにある別の施設の1日あたりの訪問者数を把握することにより、小売業者のデータサイエンティストは消費者行動のモデリングにより店舗売上の予測が可能となります。この分析

は、個々の店舗の需要パターンに応じてより適切に在庫を配分するために役立ちます。

マーケティング分野では、かなり以前から、消費者のトレンドを把握するためにサードパーティデータが利用されています。CRMシステムから取得した顧客データを、人口統計データ、購買データ、オンライン行動データでエンリッチメントすることにより、マーケティングチームは顧客の興味を把握し、自社を利用するよう顧客を促す最適な方法を特定できます。販売チームもこのエンリッチメントされたデータを活用することで、より適切に販売活動のターゲットを定め、販売サイクルを加速させることができます。

## オルタナティブデータを利用し、固有のトレンドを特定

公開会社が財務状況について公表する四半期ごとの報告書を発行すると、投資家はその情報を利用してその会社の株式を売買するかを判断します。しかし、賢い投資家は、貸借対照表はその時点の情報でしかないことを知っています。そのため、ヘッジファンドなどの金融サービス機関は、従来とは異なるソースから取得したデータであるオルタナティブデータを利用します。ソーシャルメディアのストリーム、雇用データ、衛星画像

などを組み合わせたオルタナティブデータは、上場企業が利益予測を達成できるかを、従来の財務情報が公表される前に金融機関が予測するために役立ちます。

たとえば、財務分析を強化するために、企業の店舗を訪れる消費者が増えているか減っているかを示す客足データを利用します。この情報を人口統計データと組み合わせることで、来店している消費者が会社のターゲット層と的確に一致しているかを判断することができます。匿名化されたクレジットカードの取引データ、返品データ、別の店舗の販売データを見れば、その会社について詳細に把握し、次の四半期に業績予想が上方修正されるか下方修正されるかを見極めることができます。

### イベントデータを利用し、事業への影響を把握

新型コロナウイルスの流行のような世界的な事件の影響を把握するために、あらゆる業界の企業がサードパーティデータを利用し、経営、従業員、顧客に関する意思決定を行っています。たとえば、建設会社は、新型コロナウイルスに関するデータを

利用して感染事例を特定し、感染発生場所と建設現場の位置を照らし合わせてその地域で働いている従業員数を把握し、現場を閉鎖するかどうかを判断しています。

不動産管理会社は、新型コロナウイルスに関するデータと自社の物件占有データを組み合わせ、感染症の拡大が家賃徴収にどのように影響するのかに関するインサイトを家主と共有しています。実店舗やオフィスを保有する企業は、健康データや景気回復データ（地域内の開店数や店舗ごとの売上など）を調達して事業所のデータと組み合わせ、オフィスや店舗の再開に関する意思決定をしています。

これほど役に立つ外部データを企業が活用できていないのはなぜでしょうか？その理由は、これまで外部データの利用にさまざまな課題が伴っていたことによります。



# サードパーティデータの調達は困難

外部データを調達する目的が的確であっても、それを実行する方法が効率的でないと問題となります。多くの企業にとって、以下の2つが大きな妨げとなっています。

- 従来のデータマーケットプレイスは使いづらい
- 古いファイル共有方法はエラーが発生しやすく、安全ではない

従来のデータマーケットプレイスからデータを取得するのは大変です。複数のデータマーケットプレイスからどれを選びますか？どのように始めますか？信頼できるベンダー、有用なデータ、各データの本当の価値をどのように見極めますか？

これらの疑問は、最適な外部データを特定、選択するためにはコスト、時間、労力がかかることを示しています。結局のところ、スケーラビリティに問題があるのです。各ベンダーと連絡を取り、データを評価し、取得するのをスムーズに行うことのできる方法はありません。

企業は、この時間のかかるタスクを実行するチームを作るために、膨大なコストを費やします。なかには、もともとデータモデルを構築するために雇用したデータサイエンティストにデータ取得プロセスを任せる企業もありますが、これによって本来のタスクから巨額のリソースが失われてしまいます。また、データアグリゲーターやブローカーを利用してデータマーケットプレイスでの取引を促進する企業もあります。このソリューションにはメリットがありますが、従来のファイル共有方法に依存することで生じる課題に対処することはできません。

ベンダーは引き続きFTPやAPIなどのファイルダウンロード技術を利用してデータを転送するため、多くの場合、ファイルをコピーしたり、集中的なエンジニアリング作業によってデータの抽出、変換、ロード(ETL)を行ったりする必要が生じます。開発者はさまざまなベンダーのAPIのメンテナンスやトラブルシューティングを行う必要があるため、APIさえも負担になります。また、各APIのセキュリティ手段や認証方法が異なるため、セキュリティチームに負担がかかり、潜在的なリスクが生じます。さらに、APIは大量のデータを取得するには非効率的です。

しかし、これらのデータシェアリング方法の最大の課題は、データコピーの鮮度が低く、しかもデータの取得に多大なコストがかかることです。手作業でのデータ転送では人為的なエラーが生じやすく、企業はセキュリティやコンプライアンス

に関する潜在的な問題を抱えることとなります。最悪なのは、信頼性の低いデータを押し付けられ、アナリティクスの質が低下することです。

これらの課題から、次のような疑問が生じます。時間、資金、リソースを無駄にすることなく、セキュリティやコンプライアンスを妥協せずに、大規模に外部データを利用するには、どうすればよいのでしょうか？



# データクラウドで 従来のデータシェアリングの課題を解決

外部データの調達とアクセスに、データ分析より時間がかかってはなりません。従来のデータマーケットプレイスや古いデータシェアリング手段によってもたらされる数々の障壁を取り除く必要があります。そうすることで、企業は簡単にデータソースを発見、評価し、外部データと内部データを組み合わせて迅速に分析できるようになります。

Snowflakeのデータクラウドは、データに関する課題を解決するための現代的なソリューションです。Snowflakeのプラットフォームは、マルチクラスター型共有データアーキテクチャによってすべてのデータをデータクラウド内の単一の場所に一元化し、パブリッククラウドプロバイダーや地域の垣根を越えてSnowflakeの顧客、パートナー、データプロバイダー、データサービスプロバイダーを結び付けるネットワークを形成します。その結果、データを隔てる従来の障壁が取り除かれ、データサイロが解消されます。企業はすぐに、統制された安全なデータにアクセスして企業内や企業間で共有できるようになります。

それこそが、ライブデータを元の場所に保管したまま共有しアクセスできるようにする安全なデータシェアリングをSnowflakeが実現する目的です。データコピーは一切作成されず、データの送信も行われません。データが移動しないのです。

一元管理された単一の最新バージョンのデータのみが存在します。アクセス権を付与されたユーザーは誰でも管理された安全な方法でデータを参照でき、データを物理的に取得することはありません。単一のバージョンのデータにあらゆる変更が反映されるため、複数ユーザーの同時実行による遅延や競合が生じることなく、アクセス権のあるすべてのデータコンシューマーが最新のデータにアクセスできます。

また、安全なデータシェアリング機能を利用することで、共有データを即座に既存のデータと組み合わせ、迅速に分析することができます。クエリ可能な形式のデータを取得できるため、データを複製、変換、加工する必要はありません。データ分析の遅延という問題は、過去のものとなります。

また、Snowflakeはクラウドに依存しないアーキテクチャを採用しているため、企業はクラウドインフラストラクチャー、地理的位置、クラウドプロバイダー（AWS、Azure、またはGoogle Cloud Platform）を問わず、すべての共有データに即座にアクセスできます。Snowflakeの安全なデータシェアリング機能は、Snowflakeのアカウントを所有していない企業とのデータシェアリングにも利用できる、グローバルかつ開放的な機能です。

# SNOWFLAKEデータマーケットプレイス： 安全なデータシェアリングの活用

Snowflakeの安全なデータシェアリングは、新しい現代的なデータマーケットプレイスの基盤となる技術です。

Snowflakeデータマーケットプレイスでは、データコンシューマーはクエリ可能なライブサードパーティデータにアクセスできます。Snowflakeを利用すれば、ベンダーを探したり鮮度の低いデータをダウンロードしたりして時間を無駄にすることなく、法令に準拠した安全な方法で、簡単に、滞りなく、バーチャルで、瞬時にライブ外部データを評価してアクセスすることができます。

データ変換の手間がなくなるため、外部データと既存のデータの統合が迅速かつシームレスになり、データ分析を即座に開始することができます。

また、エンリッチメントサービスを利用して自社のデータの質を高めることができます。従来の方法ではデータの複製や変換に時間がかかり、セキュリティやコンプライアンスの問題が生じる可能性もありましたが、これを行う必要がなくなります。その代わりに、Snowflakeデータマーケットプレイスで自社のデータの一部をデータプロバイダーと共有し、エンリッチメントや拡張を行った後に、自社のSnowflakeアカウントに安全かつ直接的に再共有してもらうことができます。

外部データを利用する目的が、データセットを拡張してビジネスアナリティクスを改善することであれ、AIモデルや機械学習（ML）のトレーニングに活用してデータサイエンスを強化することであれ、Snowflakeデータマーケットプレイスには次のような明確なメリットがあります。

- **データディスカバリーが容易**：クエリ可能で、内部データと統合でき、データモデリングに利用でき、BIツールに追加できる幅広いデータセットに、単一の場所から高速かつ簡単にアクセスできます。
- **最新のライブデータ**：もうデータの鮮度低下について心配する必要はありません。手作業での処理やスケジューリングが一切不要になり、サードパーティデータプロバイダーによるすべてのアップデートが即座にデータセットに反映されます。
- **コストの削減**：データのロードと変換、APIの統合と管理に伴う不要なデータアナリティクス費用を削減できます。また、データを移動させず、データへアクセスするだけなので、サードパーティデータのストレージコストは発生しません。
- **パーソナライゼーション**：自社固有のデータニーズに合わせてパーソナライズされた安全なデータフィードをリクエストすることができます。
- **内部データのエンリッチメント**：自社のデータをプロバイダーと安全に共有してエンリッチメントサービスを利用することにより、自社のデータの質を改善できます。
- **オールアクセス**：あらゆる大手クラウドプロバイダーのサードパーティデータに迅速にアクセスできます。



# 今すぐデータコンシューマーとしての能力を高めましょう

Snowflakeデータマーケットプレイスを利用し、サードパーティソースのデータに安全にアクセスして迅速にデータを統合できるようにすることで、アナリティクスを充実させ、適切なインサイトを獲得し、より情報に基づいた意思決定を行えるようになります。

Snowflakeデータマーケットプレイスが生み出す違いを、ご自身で体感してください。[snowflake.com/data-marketplace](https://snowflake.com/data-marketplace)で、無償トライアルのお申し込みを受け付けております。マーケットプレイスでデータプロバイダーを見つけて、すぐにデータのクエリを始めることができます。

残された質問はひとつだけです。あなたはデータを活用する準備はできていますか？





## SNOWFLAKEについて

Snowflakeは、Snowflakeのデータクラウドによってあらゆる組織が自らのデータを活用できるようにします。顧客企業はデータクラウドを利用してサイロ化されたデータを統合し、データを検索して安全に共有しながら、さまざまな分析ワークロードを実行しています。データやユーザーがどこに存在するかに関係なく、Snowflakeは複数のクラウドと地域にまたがり単一のデータ体験を提供します。多くの業界から何千ものお客様（2022年1月31日時点で、2021年のFortune 500社のうち241社、2021年のForbes Global 2000 (G2K) のうち488社を含む）が、Snowflakeデータクラウドを全社で幅広いビジネスに活用しています。詳しくは、[snowflake.com](https://www.snowflake.com)をご覧ください。



© 2022 Snowflake Inc. All rights reserved. Snowflake、Snowflakeのロゴ、および本書に記載されているその他すべてのSnowflakeの製品、機能、サービス名は、米国およびその他の国におけるSnowflake Inc.の登録商標または商標です。本書で言及または使用されているその他すべてのブランド名またはロゴは、識別目的のみ使用されており、各所有者の商標である可能性があります。Snowflakeが、必ずしもかかる商標所有者と関係を持ち、または出資や支援を受けているわけではありません。