

Teradata 活用事例

# ヤフー株式会社

ヤフーは、1日の処理データ量が90テラバイトを超える巨大情報サービス企業です。

より高度な分析を簡易にできる環境を構築するために、

同社はテラデータのエコシステム・アーキテクチャを採用し、

それを構成するTeradata IntelliSphere™を世界で初めて導入しました。

ヤフーのデータ戦略とそれに適合する分析基盤について、

チーフデータオフィサー 佐々木 潔 氏にお話を伺いました。

※この記事は2018年1月実施のインタビューに基づいており、役職などは当時のものです。

## データドリブンカンパニーを目指すヤフー

ヤフー株式会社は、2018年4月に就任した代表取締役CEO川邊健太郎氏の体制のもと、「データドリブンカンパニー」になることを新たな目標として掲げています。

「ヤフーは100を超えるサービスを提供しています。これらのサービスを通して収集されるデータを分析して得られるインサイトは、他社にない強みとなっています。このようなインサイトを、誰でもが素早く利用できるようにすることが、ヤフーのデータ戦略の骨子です」と執行役員 データ&サイエンスソリューション統括本部長 チーフデータオフィサー(CDO)佐々木氏は言います。

「例えば飲料メーカーは、年間に3,000アイテムくらいの新商品を出しますが、そのうち定番商品として残るのは2アイテム程度です。この状況からも、市場調査やアンケート等からユーザーニーズを正確に理解することが難しいということが分かります。何が誰にどれだけ売れているのかといったアフターマーケット情報だけ集めても、半分しか分かっていないと言えます。私たちは、消費者がある製品を買うまでに実行した検索クエリーをログとして蓄積しているの、製品購入までの行動を時系列で理解することができる。購入前と購入後のデータを繋げることで、開発・生産・購買・物流といったサプライチェーン全体に有効なインサイトを得ることができます」(佐々木氏)。

## 膨大なデータを連携しつつガバナンスを確立するにはCDOが必要

ヤフーでは2017年からCDOを設置し、佐々木氏が初代CDOに就任しました。「それまでサービス単位で所有し個別に活用していたデータを、連携して相互に利用できるようにする取り組みを2015年から始めました。システムの連携は進み、データ分析の環境は改善されましたが、全社的なデータ戦略立案や、利用ルールやガバナンスを確立するためには、全社横断的な立場でデータを統括する役職が必要だと分かりました」(佐々木氏)。ヤフーがCDOを設置した背景には、このような事情があったのです。

## 分析者にとって最適なデータ環境を求めて

ヤフーのデータ戦略に最適なデータ環境とはどのようなものなのでしょうか。データサイエンティストやエンジニアだけでなく、社員全員が日々データに基づいたビジネス判断をするようになることが、データドリブンカンパニーとして目指す姿だと佐々木氏は言います。「Webマーケティングにおいて離脱率や滞在時間は相変わらず重要な指標です。ちょっとしたレイアウト変更やフォントの違いで、これらの値が変化します。そこで画面を少し変えてはテストするわけですが、デザイナーが自らクエリーを発行し

# YAHOO! JAPAN

### ヤフー株式会社

ポータルサイトYahoo! JAPANをはじめとする、100を超えるオンラインサービスを展開する巨大情報サービス企業。

URL: <https://www.yahoo.co.jp/>  
設立: 1996年

### 課題

・複数のデータソースに存在するデータを利用した分析の高度化

### ソリューション

・Teradata IntelliSphere™  
・テラデータのエコシステム・アーキテクチャ

### 導入効果

・適材適所のデータソースの連携と分析の高度化  
・ソフトウェアのライセンス管理と予算管理の簡易化

て、結果を分析できれば生産性が飛躍的に高まります」と、データ&サイエンスソリューション統括本部 データプラットフォーム本部 遠藤氏は言います。

全社員がデータ分析を日常的にできる環境を実現するには、必要なデータがどこにあるのかをユーザーが意識せずともアクセスができるようにすることが重要です。そして、ユーザーが探索的な分析、つまり自由検索ができる柔軟性と、千単位のユーザーが同時にアクセスしても、クエリーのレスポンスが速いことは必須といえます。

## サービス毎にバラバラに存在するデータを繋げるための考え方

ヤフーが求めるデータ環境を構築するためには、どのような考え方が必要でしょうか。「データソースを1つに統合するのは現実的ではありません。ヤフーにも既にTeradata Database、Hadoop、Cassandraなど多様なデータソースが存在します。これらをどれか1つに統合するのは技術的にも難しく、コストもかかります。そもそも適材適所に導入していますので、統合するのはむしろ弊害があります」(佐々木氏)。しかし、データ分析の価値を上げるためには、「これらのデータを有機的に繋げる必要はある」と佐々木氏は強調します。適材適所に配置されたデータソースを有機的につなぐことで、必要なデータをそのデータソースを意識することなく分析することができる—これがまさしくテラデータの考えるエコシステム・アーキテクチャです。「テラデータ

が提案するエコシステム・アーキテクチャを見た時に、ヤフーにとっての理想的なアーキテクチャとして描いていたものと基本的なコンセプトが同じだった」とデータ&サイエンスソリューション統括本部 データプラットフォーム本部 櫻井氏は言います。

## ヤフーのニーズに適合したTeradata IntelliSphere™

テラデータのエコシステム・アーキテクチャを構築するために必要となる、「データ収集」、「アクセス」、「デプロイ」、「管理」の主要機能を包括したソフトウェア・ポートフォリオが、Teradata IntelliSphere™です。

Teradata IntelliSphereは、現在9種類のソフトウェア・コンポーネントで構成されており、それらを自由に利用できます。サブスクリプションベースのライセンスで提供されており、常に最新バージョンの製品を使うことができるほか、将来的に新しいソフトウェアがTeradata IntelliSphereの一部としてリリースされた場合は、その機能をライセンス価格はそのまま利用することができます。

ヤフーでは、Teradata QueryGridを利用して、ヤフーが導入している複数のTeradata Databaseを高速に連携すると共に、Hadoopと他のシステムの連携に活用しています。ヤフーの実装ロードマップにはHadoopやその周辺システムについて、オープンソースのクエリーエンジンであるPrestoとTeradata QueryGridを使った連携が含まれており、将来

的にHadoopやPrestoの接続先システムの台数が増えても金額が変わらないという、Teradata IntelliSphereの柔軟なライセンス体系についても、ヤフーは高く評価しています。

「データ分析は小さく始めて、徐々に拡大していくため、スケーラビリティはとても重要な要素です。どのくらいデータ量が増えるかを当初に予測するのは難しいので、接続先システムの規模が大きくなっても料金がかわらないというIntelliSphereであれば、上限が明確なので予算も立てやすいのです」(佐々木氏)

## 技術会社であり続けたいからテラデータを選んだ

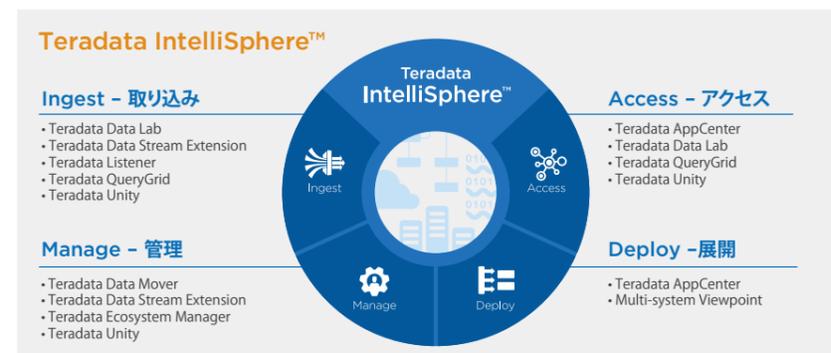
ヤフーには、技術に強い会社であり続けたいというコンセプトがあります。「自社の技術レベルを保持しつつ、自社とは違う強みを持つテクノロジー・カンパニーとパートナーシップを結んで協業ができる企業でありたい」と佐々木氏は言います。

このヤフーの想いに応えられる企業の1つとして選択されたのが、テラデータでした。ヤフーとテラデータは共同開発のパートナーとして技術提携しています。テラデータが新しいサービスやソリューションをリリースする際には、そのプラン段階からヤフーと伴に議論を重ね、ソリューションの完成度を高める努力をしています。ヤフーが世界で初めてTeradata IntelliSphereを導入した理由も、こうした技術提携により、ヤ

フーのニーズを反映したからでした。佐々木氏はさらにテラデータの姿勢も評価しています。「テラデータのことは、データウェアハウスのトップランナーと認識しており、ヤフーでもかなり昔からその製品を利用してきました。しかしここ数年の取り組みを見ていると、アナリティクスのリーディングカンパニーを目指すべく、積極的な買収やソリューションの拡大を実施し、ビジネス変革に取り組み続けています。ヤフーも変わり続けなければいけないという刺激をもらっています」(佐々木氏)。

## 日本全体が一丸となって戦えるプラットフォームを

ヤフーの掲げるデータ戦略は社内でのデータ活用の促進に留まらず、データ分析から得られたインサイトを外部の企業や自治体、研究機関などにも提供することを視野に入れています。この戦略が具現化したものがヤフーの新プロジェクト「DATA FOREST」であり、2019年の事業化を目指して始動しました。他の企業や組織にも、インサイトを提供することについて、佐々木氏はこう語ります。「GoogleやAmazonといったグローバル企業は、もはや日本企業が単独では対抗できないだけのデータ資産を持っています。しかし、日本企業が会社の枠を越えてデータとインサイトを共有し、一丸となって戦えば十分対抗できます。ヤフーは日本企業が共存共栄するためのプラットフォームを提供したいと考えています」



柔軟性の高い分析エコシステムを構築するために必要となる主要機能を包括する9種類のソフトウェアで構成される(2018年5月現在)



ヤフー株式会社 執行役員  
データ&サイエンスソリューション統括本部長 チーフデータオフィサー  
佐々木 潔 氏(中央)

ヤフー株式会社 データ&サイエンスソリューション統括本部  
データプラットフォーム本部 部長  
遠藤 禎士 氏(右)

ヤフー株式会社 データ&サイエンスソリューション統括本部  
データプラットフォーム本部 データウェアハウスチーム  
櫻井 史彦 氏(左)