解説

ITベンダーが提供するビッグデータ事例[†]

岩崎 将之^{*}

1. はじめに

前回の寄稿[1]では、IT業界のみならず企業を中心に社会を席巻している「ビッグデータ」というキーワードに関して、これにまつわるデータの特徴とテクノロジーについて紹介しました。今回はよりビジネスに即した事例を紹介します。

2. 前回の振り返り

実社会での取組みを紹介するに当たり、まずは前回のポイントを振り返ります。

ビッグデータとは、3つのVで表わされる特徴があります。

1. Volume：ストレージコストの低下や実社会のあらゆる局面での「トランザクション」が高精細化・多頻度化することにより、データの量が増えてきています。
2. Variety：センサーや携帯電話などデバイスの状態をデータとして捉えられるようになる動きや、facebookやTwitterなどのソーシャルネットワークのデータをビジネスに生かす動きなどからデータの種類が多様化してきています。
3. Velocity：デバイスの進化や多様化・増加により、同時に多くのデータを捌く必要性が高まってきています。

ビッグデータを支えるテクノロジーは広範になってきています。（図1）

1. データベース：従来より多くの業務部門のユーザーや高度な分析をするユーザーが「同時に」、
「沢山の」問い合わせをするだけでなく、自動化された仕組みからの問い合わせにも答える基盤

として、これまで以上にデータベースの重要度は高まっています。（図1ではBig Data Managementの領域。）

2. Hadoop：多様なデータの中にはそのままでは分析しづらい文章のようなものもあり、こうしたデータを分析できる形に「高速に」「加工する」用途や、将来分析したくなる可能性に対応するために「保存しておく」基盤として、昨今注目を集めています。（図1ではBig Data Managementの領域。）
3. 統計解析：単に表やグラフを表示するだけの従来のデータ活用だけではなく、高度な統計解析をするニーズに答えるためのツールの領域において、R言語のようなオープンソースが注目を集めています。（図1ではBusiness Analyticの領域。）
4. ミドルウェア：前回は詳しく触れませんでした。が、センサーやデバイスが発達すると、大量なデータを即座に取り込み、更に即座にアクションをする必要性が高まります。工場の生産機器から届くデータから、不具合を示すデータを受け取った際に即アラートを上げるといった例があります。（図1ではFast Data Processingの領域。）

3. ビッグデータ活用事例1、消費者との距離を縮めたい製造業

製造業はこれまで、実際に商品を購入・利用する消費者の情報に直接触れる機会が少ない業種の一つでした。製造業の中でも一般消費者がその商品を利用する一般消費財メーカーなどは特に、消費者の声や購買行動の情報を得たいとしながらも、従来の消費者との接点は、商品に不具合が発生した時に問い合わせをするサポート部門が代表的です。そこでは主にネガティブな情報が中心に寄せられていますが、その中からも「よりよい商品を作るためのヒント」を見つけ出して新商品の開発に生かしたり、真摯な対応をすることで顧

[†] The introduction of “Big Data” in practice
Masayuki IWASAKI

^{*} 日本オラクル株式会社 製品戦略事業統括本部 戦略ビジネス推進部
Product Strategy Business Unit/Emerging Business Development/Oracle Corporation Japan

図1 オラクルがお手伝いできること：あらゆるデータを包括的且つ迅速に活用する為の技術を提供

図2 2012 Oracle Innovation Award Winner お客様事例
イノベーティブな顧客体験を提供する斬新なビジネスモデル

ナイキ 様

- Nike+ DIGITAL SPORT を支えるシステムプラットフォーム
- 【ご採用製品】 Oracle Coherence, Oracle Exadata & Oracle Spatial
- 800万人のグローバルユーザを対象
- 日々の生活環境におけるライフスタイルをモニタリング

ORACLE

Copyright © 2013, Oracle and/or its affiliates. All rights reserved.

図2 2012 Oracle Innovation Award Winner お客様事例
イノベーティブな顧客体験を提供する斬新なビジネスモデル

客満足度を向上させようと取組んできました。

ここに、ビッグデータは新しい情報をもたらしています。米 Nike 社は、消費者が自らの腕につけて活動量を測定する「Nike+FuelBand」を提供しています(図2) [2]。この活動量は独自のNikeというブランドから想

起されるランニングやスポーツでの活動だけではなく、睡眠時間や歩数など日常生活における行動からも活動状況を算出し蓄積することができます。活動量は独自の単位「Fuel」としてバンドに表示されるだけではなく、スマートフォンやPCで確認ができるようになっ

ており、個々人が設定した目標の活動量に対しての達成度と現状を比べることで、運動に対するモチベーションを高める効果があります。さらに、独自に準備されているソーシャルネットワーキングサービスを介して、利用者同士が相互に情報を交換することもできます。この利用者は800万人を越え、これを支えるシステムには、随時リアルタイムで膨大な量のデータが入ってくるだけではなく、システムから上述のような活動の実績や達成度をフィードしています。システムの特性上重要なのは、「同時に」「瞬時に」データのやり取りを捌く仕組みと、「大量の」データを蓄えておき、しかも後々閲覧する際に「ストレスなく」素早く提供する仕組みです。前者のポイントを支えるのは、インメモリで大量のデータを受け取り、かつユーザーに返すことができる製品であり、図2に表記されているOracle Coherenceがこれにあたります。後者のポイントを支えるのはOracle Exadataであり、これはOracle DatabaseというRDBMSを最適に安全に運用するためにオラクルが開発した、ハードウェアとソフトウェアを組み合わせた製品です。また、Oracle Spatialという製品は、位置情報を扱うことができる、Oracle Databaseのオプション製品です。

これにより企業はどんなメリットを享受することができるのでしょうか？それは、企業にとって、「消費者との接点が増える」ことです。従来、スポーツにまつわる製品を購入するのは、何かを始める時、何かが摩耗して新しく買い替える時、新商品のプロモーション

をテレビや雑誌などで見た時、などでした。これでは、競合企業・商品との差別化は難しいと気付いたのです。少ない消費者との接点において、如何にブランドイメージを高めていくか、店舗で良い場所に商品を配置してもらうかという以外に、「いつも自社の商品やブランドに接してもらい『消費者と常に接してもらう』ことで、別のブランドへの離反を防ぐこと」、さらには「このサービスを目的に自社の商品を買ってもらうこと」で競争力を高めていけるのです。

4. ビッグデータ活用事例2、 大量のデータから新発見を突き止める 研究機関

アメリカ国立ガン研究所では、大量の「非構造化」データを分析することにより、社会に貢献する新たな発見を支えています。5つの主要なガンを表す単語と、ガンに関係する17,000の遺伝子を表す単語の相関関係を分析するために、2,000万もの医療関連の出版物の内容を総ざらいしています。医療関連の文献の中に出てくるガンの単語と遺伝子の単語が出てくる量や組合せを総点検します。このためには、テキスト化された全ての文献にキーワードが出てくる組合せを確認せねばならず、大量の「キーワード同士のマッチング」処理が必要です。ここに、Hadoopが活用されています。Hadoopが得意な並列分散処理により、従来处理しきれなかったり時間がかかり過ぎていた分析作業が短時間で結果に辿りつくようになりました。（図3）

図3 アメリカ国立ガン研究所
遺伝子とガンの関係性を特定

解決策

- 17,000の遺伝子と、5つの主要なガンの関連性を、2,000万の医療関連の出版物のアブストラクトを通して関連性を算出
- 6,000万の患者の遺伝子と9億のシミュレートされたmiRNAの関係性を算出
- これらの遺伝子が作用する経路および、これらの遺伝子をターゲットとする薬に関しての知見を深めることにより、研究者が自身の研究を進める手助けをする

“Our mouths were open in amazement”

Dr. Robert Stephens
Director of Bioinformatics,
National Cancer Institute

The winner of the 2012 Government Big Data Solutions Award is the NCI Frederick National Laboratory

“This will help researchers accelerate their work in areas of importance for all humanity...The promising approach of this solution is repeatable across many other Big Data challenges for bioinformatics...”

- CTOVision.com

* <http://ctovision.com/2012/11/the-winner-of-the-2012-government-big-data-solutions-award-is-the-national-cancer-institute/>

ORACLE

Copyright © 2013, Oracle and/or its affiliates. All rights reserved.

図3 アメリカ国立ガン研究所 遺伝子とガンの関係性を特定

研究文献の総量から大まかな傾向を導き出すことで今後の研究の方向性を見定めるといった発想は、それ自体は効果的な結論に直結しにくいかもしれませんが、医療研究ひいては社会全体の効率性の観点から見ると画期的です。ともすると何代にも渡ってそれぞれの研究機関が時間を掛けて試行錯誤するような時間が、大幅に短縮できるかもしれません。少なくとも、将来振り返ってみると間違った方向に時間を費やすという非効率性を減らすことが期待できます。このような大量のデータを高速で何度も何度も参照するような処理能力は、これまで限られた機関や企業しか持ち合わせておらず、費用がかかっていたものです。しかも、本来研究に集中したい研究者の方々が、ITのスキルを得るために本来とは異なる二次的な努力を求められていたともいえます。この事例では、オラクルのOracle Big Data Appliance というHadoop 用のシステムが活用されています。研究者や企業の担当者は、新しい処理を動作させるための言語を覚える必要があるものの、ハードウェアやネットワーク、そしてOSやHadoop の設定自体の知識や構築作業からは解放されるため、本来の研究や業務に集中できます。さらに、Hadoop を動作させるための言語自体にもITベンダーは「より使いやすい」仕組みを提供しており、例えばデータベースで分析するための言語として長年の実績と普及を遂げているSQLでHadoop の処理をすることも可能になってきています。

5. おわりに

以上、二回にわたりITベンダーの視点で、ビッグデータの定義や特徴と、事例をご紹介してきました。ここでお伝えしたかったのは、企業のみならず社会に存在するあらゆるデータを蓄積して分析することは、企業活動や社会に関する研究活動を新しい段階へ進める原動力となっているということです。従来IT業界がお伝えする事例は「速い」「見やすい」「安い」などの観点で語られるものがその多くを占めていましたが、よりビジネスや社会貢献の視点で語られるようになってきているのです。こうした視点でITベンダーがお客様や社会を捉えて、そこに貢献できる製品を提供するようになったことが、ビッグデータというトレンドがもたらした変化のひとつであると、筆者は考えております。

参 考 文 献

- [1] 岩崎将之, ITベンダーのビジネス開発担当者から見たビッグデータ, 知能と情報(日本知能情報ファジィ学会誌), Vol.25, No.5, pp.153-157, 2013.
- [2] Nike + FuelBand SE, http://www.nike.com/us/en_us/c/nikeplus-fuelband, (参照2013.11.30)
(2013年12月1日 受付)

[問い合わせ先]

〒107-0061 東京都港区北青山2-5-8

オラクル青山センター

日本オラクル株式会社 製品戦略事業統括本部 戦略ビジネス推進部

岩崎 将之

TEL: 03-6834-3269

E-mail: masayuki.iwasaki@oracle.com

著者紹介

いわさき まさひろ
岩崎 将之 [非会員]

2000年慶応義塾大学環境情報学部卒業。米系ITベンダーにてデータベース事業部に所属。金融・製造・流通の顧客への情報系データベースシステム構築コンサルティング、プロジェクトマネジメントならびに、データを活用した組織変革・業務変革コンサルティングに従事。2011年より日本オラクルに在職。