

## ソーシャルビッグデータの現状と今後



特 集

### ●会社の概要

本日は「ソーシャルビッグデータの現状と今後」について話します。まず当社の概要です。オフィスは東京・渋谷にあります。役員は取締役6名、監査役3名。社員数は約60名。資本金は3億7000万円程度。勤務時間ですが、午前9時から午後6時まで、フレックスタイム制です。IT企業は仕事時間が関係ないケースが多いのですが、当社は年末年始、夏季休暇も必ずとっています。私も先日、9連休でベトナムに行ってきました。社員は大体午前10時に出社し、午後8時には帰ります。社訓の中で「公私とも大切に」を掲げています。取引先は博報堂、富士通、NEC、TBS等、上場している大企業がほとんどです。私の座右の銘は「三方よし」で、売り手よし、買い手よし、世間よし。この中で「世間よし」は世の中をよくするビジネスモデルを提供したいということです。ビッグデータに関しては、ビッグデータを取り扱って世の中をよくしていきたいと考えています。

### ●新聞掲載

今年6月30日、7月1日の2日連続で新聞に記事が載りました。このようにアジア版と、ベンチャー企業欄と2日連続で掲載されたのはとても珍しいようです。どんなことかという、例えばパソコンの情報データとして、災害などが発生したときに最も早く流れるのはツイッター。その情報を集めることで、いち早く今何がおこっているかが分かるサービスをつくった。当社はベトナムに会社を持っていますが、ベトナム市場の状況はなかなか分かりません。東南アジアのアンケート調査のパネルもあまり意味がないし、そうした場合にソーシャルでどのように言われているのかによって、自社商品の意外な使われ方があるということです。

4月17日の日経新聞、これは1面トップ記事になりましたが、ビッグデータの300社と連携したコンソーシアムを立ち上げることが記事で取り上げら

データセクション株式会社 代表 取締役社長  
澤 博 史 氏

れた。いろんな人がデータを持っていますが、何のデータがあるのかが分からない。それでは何も進まないわけです。だから簡単に言うと、データとして何があるのかの一覧を整理し、可能ならそれをデータベースに保管して、誰もが一部を使えるようなサービス基盤をつくりましょうというものです。その考え方に賛同していただき、3年後には300社を集めようとしています。最初に30社を集めるつもりでしたが、すでに90社が集まっていたので、現在も運営をしています。

4月2日にはTBSと業務資本提携をしました。テレビ業界もテレビ離れに危機感を感じています。視聴率の評価だけでなく、ソーシャルメディアで番組の中身がどれだけ評価されているのかをアピールしたい。ソーシャルメディアのデータを分析して、他のサービスに比べてこれだけ効果があるとアピールしたいという思いがあります。

### ●自己紹介

ここで簡単に私の自己紹介をします。大阪市立大学理学部を卒業し、富士通に入社。はじめはC言語のプログラマー、その後に事業企画、そしてマー



講師 澤 博史氏

ケティング戦略立案部門からアウトソーシング企画部門へ。その頃にブログがブームになり、私も興味がわきました。富士通をやめる平成16年ころですが、ある日本でいちばん売れている石鹸がありました。この石鹸実は、当初はニッチな石鹸でしたが、あるとき東京のOLが現地で買って、帰った後にブログで絶賛しました。友達も買うという連鎖がおこっていった、この石鹸がある販売店での売上ナンバー1になりました。たぶん今でもそうだと思います。この話を10年くらい前に聞いて、今後は商品自体もニッチになっていくのではないかと思いました。

「この石鹸はいいよ」という流れをみると、次ほどの石鹸が売れるかが分かるのではないのか。当時に私が思ったのは、次に売れそうな商品を買占めて商社をやろうと。国内で売れるものを海外に持っていき、東南アジアで売れるものを日本に持ってくる。そんな商社をつくらうと思ったのですが、富士通をやめてから双日に入社しました。ここではファイナンスや新規事業企画などを勉強しました。商社をつくるにしても、投資を勉強する必要があるということから、投資の会社イーライセンス社に投資。そうこうするうちにデータセクションの前社長からの要請もあり、平成21年7月にMBOする形で私が社長に就任し、新たにスタートしました。

### ●マクロ的環境および目指す方向性

素材によって、社会のインフラになるようなイノベーションをおこそう。それが将来の目指す姿です。現状としてソーシャルリスニング、何を書いているかを見る、分析するにとどまっていますが、将来的にはCRM（顧客関係管理）や人材育成に使うとか、リアルなインターネット、テレビCMの効果分析に使われたりするのではないか。将来の商売として、こんなことがあるのではないのでしょうか。

### ●データセクションのポートフォリオ

これは当社のポートフォリオですが、当社のサービスには3つの軸があります。まずはソーシャルメディアの情報を1日あたり数千万件集めていて、すでに12年になるからかなりのデータがあります。そのデータを基に何を書いているかを分析するサービス、つまりそのツールを提供しています。これらのツールが当社の売り上げの半数以上を占めていま

す。何を書いているかが分からないので分析してほしいと言われる。そのコンサルティングもやっています。コンサルをしていくと、何かつくってほしいと言われますから、その開発もしています。この3つのポートフォリオで事業を進めています。分かりにくいと思いますが、せいぜいこれで100億円くらいの市場。しかしビッグデータはそんなところに市場規模があるのではないと思っています。これは金のなる木で、次の事業を生み出すための柱となっていくもので、試行的にやることでこれが広告の効果測定にもつながっていきます。

### ●今のデータセクションとは

データを集めながら、どういうことが強みなのかについて説明します。データを集め、それをソース(SaaS)のツールで提供し、コンサルティングをしていると、こういうものをつくってくれと要望が出てくる。運用を進めていると、このデータも含めて分析してほしいと言われる。ビッグデータの中でいちばん大事なのがPDCAを早く回すことなのです。いくら1年をかけて計画してやってみたとところで変わってしまう。だから2~3週間、1カ月、2カ月のスパンで結果を出していき、その結果が正しいかどうかを検証。それを回していくことで最適値を求めていくのが大事なことです。そのためにはデータを持っていないとできませんし、コンサルも必要ですし、システム開発の体制も必要です。こうした体制が整っているのは当社だけだと思います。例えばソーシャルメディアのデータの5年前からのものを今から集めようとしても無理なことです。せめて2~3年データを集めないと事業になりません。競合するにも、2~3年後にしか事業が始まりません。

### ●ツイッターから見る「幸せ感じる!」ランキング

ソーシャルメディアでいったい何ができるかを具体的に見てみましょう。SaaSのエンジンによってツイッターで「幸せ 感じる」で検索すると何万件、何十万件が出てきます。そのデータについて品詞分解を行って、ランキングを付けます。その結果がこれです。「幸せ感じる!」ランキングの1位は「イケメン」なのです。こうしたことが分かる。分からないことが分かるのがソーシャルメディアです。



### ● 「ラーメン食べたい！」のはいつ？

同じくソーシャルメディアで「食べたいもの」と検索し、分析して名詞ランキングにすると、1位がラーメン、2位が肉、3位がアイスというように出ます。そこで「肉 食べたい！」とソーシャルのデータを分割すると、ピークのデータが出る。ピークには特徴があって、じつは「29（ニク）の日」なのです。29の日になると肉を食べたくなる人が多くなる。これはまだプロモーションされていないので、もったいないことだと思います。「ラーメン食べたい！」でやってみると、これにも特徴があります。アニメの放送で「ラーメンが食べたい」と。じつはラーメンを食べるシーンがあるのです。テレビの画像を見てラーメンを食べたいと思う人が増えてくるのです。例えばこのアニメの画像で「ラーメンを食べたい」と思うことがリアルタイムで分かれば、これが終わった後にラーメンを販売されている事業者さんがラーメンのCMを出せば、5倍、10倍、100倍もの効果がありそうです。そういうことをしっかり分析することで、広告の価値をもっと高めることができるのではないのでしょうか。

### ● ソーシャルビッグデータを活用した現状の分析例

これからはソーシャルビッグデータを活用した分析例を紹介します。

#### ●（事例1）人気タレントの比較

人気タレントのソーシャルでの分析レポートを定期的に出すことをやっています。その例ですが、月ごとにトータルで人気度が上がっている、下がっているということが話題になっています。男女のどち

らに多くつぶやかれているのかということ、ある芸能人さんは男性側、また違う芸能人さんは女性側でのつぶやきの数値が多くなっています。これもじつは機会学習であって、この単語を使っているのは男に違いない、女に違いない。若者ごとに使う単語が推測できて、このつぶやきは男の人、20代の男、30代の男だと出すことができます。それによって分析したのがこの結果です。当社で200人くらいを抽出してテストをしたところ、80%くらいの精度がありました。例えば30年前のアイドルのことを、まず10代ではつぶやかかない。新入社員に先日聞いてみたら、30年前のアイドルを知りませんでした。例えば学生の場合、我々の世代なら中間試験の時につぶやいた。そういう言葉をつぶやくかどうかによって年代や性別が分かるといったようなことです。

他に形容詞ランキングでは、ある芸能人さんは「やばい」、他の芸能人さんは「おもしろい」、また他の芸能人さんは「美しい」、また他の芸能人さんは「すごい」、また他の芸能人さんは「かわいい」と言われます。これをグラフで表すと、このように完全に分かれてしまいます。皆さんはあまりご存じないようですが、このようなことができます。ちなみにこれをビールにあてはめて、キレ、コク、おいしいなどでやると、そのビールがどう飲まれているかの相対比較ができるわけです。うちのビールはキレがないから、少しキレをアピールしようとプロモーションをする。その後でキレがどのように変化していくが見えたりします。このように定量的に見えたりして変わっていきます。

#### ●（事例2）ビールの分析例

もう1つ、分析レポートのビールの事例ですが、接触、体験、ファン化という分け方をよくしています。商品名がつぶやかれることは売れているということです。買った、飲んだ、おいしかった、まずかったというような、経験の言葉を辞書化するのは。そうするとその人たちは体験したということになります。また買った、いつも買った、今回もまた、こういうことを辞書化することで、ロイヤルカスタマーになる。それによって数を分析してみると、接触している人何人に対して、体験が何人かが分かるわけです。

### ●（事例3）地図情報とソーシャル情報を合わせた活用例

地図情報とソーシャルの情報を合わせた活用例として、例えば「混雑」「すごい人」「混んでいる」「行列」「並んでいる」「事故」などをキーワードにします。そうすると、渋谷で最近できたスポット「ピカリエ」で、ピカリエ「すごい人」「すごい人」「すごい人」とつぶやかれたとすると、このキーワード「すごい人」という、人が集まっている所を地図上に示すことができます。どこにどれくらいの人が集まっているのか、例えば野球が終わった、相撲が終わった、結婚式が終わった、電車が事故で止まっている。そうしたキーワードでGPS探知機をつくる。タクシーは駅で待っていて、客を乗せて送っていき、空車で帰ってくれば実車率50%。帰りに乗せて帰ることができるかどうか勝負です。だから、帰りに人が集まっている所に回って乗せて帰ってくることができるということで、タクシー業界で使われ始めています。

### ●（事例4）画像を利用した即時に情報共有できる仕組みの事例

火事、事故、災害などをキーワードにすることによって、何が起きているかをいち早く画像でとらえることもできます。それがこの事例です。こういうのを辞書にすることで、ソーシャル上で今危ないと言われる情報を位置情報と合わせて見ることで、どの地域で何が起きているかが、ほぼリアルタイムで検知することができるわけです。大阪駅で飛び降り自殺をしようとしている人がいたとして、それを皆さん全員が見られた時にたぶん警察に通報しないと思います。しかし、ツイッターで自分の周囲に対して通報する人がいると思います。警察への通報より、ソーシャルで上がってくる方が早いのではないのか。それを利用したシステムをテレビ局さんに導入。全社員で使っていただき、全員のパソコンで何があったかを再現的に見ることができるようになっていきます。

### ●流通小売でのバイトテロ対策での利用例

次は、流通小売でのバイトテロ対策の利用例です。最近のコンビニではアルバイトが冷蔵庫の中に入ったとか、いろんな問題事例があります。企業の皆さ

んには、防止するためにいちばん大事なのは、ソーシャルで「監視カメラがあります」という状態をつくることです。例えば店の裏に各店舗の状況が流れていると、監視されているからそんなことはできないと思う。そのような抑止力を考えた方がよいわけで、実際に不正投稿が上がってきたら、リアルタイムで店舗に通知する。こうしたことをサービス化していて、流通業界では導入していただいています。

### ●ベトナムでの分析事例

ベトナムでの事例を簡単に紹介します。当社はベトナムと東南アジアのデータを持っています。ある即席麺が売れているようです。その即席麺がどんな理由で売れているのかを名詞ランキングで見ると、「おいしい」「酸っぱい」「安い」「風味がいい」「未調理がおいしい」と言われている。「未調理がおいしい」とはどういうことかと分析してみると、「そのまま食べても、お湯に入れても炒めてもおいしい」と言われているのが分かります。日本でインスタントラーメンをそのまま食べることは聞きませんが、ベトナムでは10代、20代の3割くらいはそのまま食べるようです。日本では考えられないが、ベトナムではこういうことが起きているというヒラメキ眼というのが、ソーシャルメディアでの特徴でもあります。

### ●ソーシャルビッグデータを活用した今後のビジネスモデル例

#### ●（事例A）会員サイトでのレコメンデーション向上（ソーシャルCRM）

会員サイトで何を見ていて、それに関心があることが分かれば、メールを出せるし、ホームページ上でバナーとか広告にアピールすることもできる。こうしたレコメンデーションに使われるやり方が最近増えています。

### ●ソーシャルCRMビジネス

既存の接触情報、つまり来店している人たち、購買情報、サイトアクセスの情報に加えて、何が好きなのか、どのテレビを見ているのかが分かる。それによってレコメンドの効率化ができるのではないかという取り組みが最近行われ始めています。

### ●ソーシャルメディアから株式予測ポートフォリオ作成

1つの事例として、ソーシャルメディアで株価が予測できるという取り組みをゴールドマン・サックスなどでは始めています。例えばアナリストの中から成果の高い人を10人選び、10人が10銘柄、合わせて100銘柄を会社名とそのブランドで分析し、件数が少ないものはカットして、件数の多い企業でポジティブの割合が多い企業だけをとってきて、ベスト10をつくる。それだけポートフォリオをつくると、机上ではアナリストより運用実績がでます。これによってファンドをつくって、技術によってこのファンドの成績が良くなったり、悪くなったりする時代が来るのではないかと思います。これを頭のいい方がつくることで、株だけで1~2年くらいはかなり儲けられるのではないかと私は思っています。ただ私はこのビジネスモデルはあまり好きではなく、このエンジンを全員が真似してしまったら、世の中で皆が真似してしまい、このエンジンの意味がなくなってしまいます。世の中のためにならない可能性があるため、私はあまりやりたくありません。

### ●データエクステンジコンソーシアム設立の理由

最後に、先ほど触れたデータエクステンジコンソーシアムの取り組みについて紹介します。課題としては、国内で①企業が委縮する環境、②ビッグデータに興味のある人・関係者がどこにいないのか分からない環境、③ビッグデータにどのような利用方法があるのかを共有できていない環境であること。欧米ではかなり進んでいるのですが、日本だけが取り残されかねない環境にありました。こうしたことからビッグデータ・コンソーシアムを立ち上げました。今では流通大手や広告代理店など、競合を含めて加入しています。現在までに12の分科会が発足し、それぞれ具体的な事業モデルや課題を意見書の形で出せる環境が整っています。私は経済産業省主催のデータドリブン協議会のコアなメンバーである準備会合委員として参画していて、委員としてこのコンソーシアムからもサポートしていこうと考えており、法整備やガイドライン策定へと向かっていければと考えております。

