

# ビッグデータ・オープンデータ活用セミナー 新たな価値創造を

総務省北海道総合通信局情報通信部  
情報通信振興課

※1 ブロードバンド環境  
アナログ回線（DSLを除く）やISDNデジタル回線を利用したインターネット接続とは異なり、映像や音楽などの大容量コンテンツを快適に送受信できる「高速のインターネット接続環境」のこと。

※2 ビッグデータ  
インターネットの普及とIT技術の進化により生まれた、より大容量かつ多様なデータを扱う新たな仕組み。その特性は量、頻度（更新速度）、多様性（種類）により表される。

ブロードバンド環境<sup>※1</sup>の整備やスマートフォンに代表される情報端末の急速な普及により、多様で膨大なデジタルデータがネットワーク上で生成・流通・蓄積される時代を迎え、これらビッグデータ<sup>※2</sup>を活用した新産業・新サービス創出が注目されています。

また、国や市町村等が保有するさまざまな公共データを二次利用可能な形で公開するオープンデータ<sup>※3</sup>の取り組みも始まっています。

これら「ビッグデータ・オープンデータ」の利活用の取り組みについて広く理解を深めていただくために、北海道総合通信局と北海道テレコム懇談会は、2014年5月27日、札幌市において、「ビッグデータ・オープンデータ利活用推進セミナー」を開催しました。

## 基調講演

### ビッグデータ活用でイノベーションを起こすには ～Nice Try!の推進を～

ビッグデータが注目される理由の一つに、さまざまな事象を「見える化」することがあげられます。「見える化」することにより、今まで気付かなかったことに人々が気付くようになるのです。それで、意思決定や行動が変わります。一人ひとりの変化は小さくても、積み重ねることによって大きな変化が生まれます。まさに、データ活用でこのような変化が起こっています。この変化が、社会変革、イノベーションにつながるのです。



稲田 修一 氏  
東京大学先端科学技術  
研究センター特任教授

まず農業を例に示しましょう。オランダのトマトの単位面積当たりの収穫量は、日本の8.2倍です。きゅうりは13.1倍、なすは14.1倍です（いずれも2011年）。トマトの単位面積当たりの収穫量の変化を調べると、オランダでは1971年から2011年までの40年間で、5倍になっています。その間、日本は1.37倍です。

このような差がついた理由ですが、オランダでは、まずマーケティングの観点から力を入れる作物を決め

※3 オープンデータ  
インターネットなどを通じて誰でも自由に入手し、利用できるデータの総称。政府、自治体、企業などが公開する統計資料、文献などを指し、図面、動画なども含む。

ています。トマト、きゅうり、パブリカ等マーケット規模が大きな作物を選択しています。競争相手はスペインやイタリア、それからトルコ等です。土地が狭く、気象条件が良くないオランダは、競争上不利です。それで、これらの国に勝つために、より良い生育条件をデータ分析で割り出し、その条件を大規模な温室の活用と生育環境のコンピュータ制御で実現し、さらに生育条件の最適化に向け改善を続けることで、単位面積当たりの収穫量を増やしたのです。まさに、データ活用で生育条件の最適化というイノベーションを起こしたのです。

日本でもデータを上手に活用している地域があります。例えば、北海道の芽室町では、衛星のリモートセンシングデータ<sup>※4</sup>を使って、小麦の収穫時期を最適化しています。データ分析により、それぞれの畑の収穫時期を決めるのです。ポイントはタブレット端末の利用です。現場で分析結果と畑の状態を見比べ、確認することにより納得が得られるのです。データ活用により小麦の状態が良いタイミングで収穫することにより、小麦の品質アップにつながっているそうです。

今まで、コンピュータと現場はなかなかつながらなかったのですが、現場で使える端末の出現によりこれが変わり始めています。「見える化」により「なるほど」と納得すると、人の行動が変わります。判断が変わります。これがデータ活用の本当の価値なのです。

医療分野でも同じようなことが起こっています。現在は、遺伝子の分析と医療データの集積で、さまざまな病気のリスクが分かります。そのリスクを下げるための対策も次第に分かってきました。現在、お医者さんの役割は、「病気になったら治す」ことですが、未来のお医者さんの役割は、「病気にならないように予防する」ことにも変わるかもしれません。データを集積し、人の健康を「見える化」することにより、医療でも大きなイノベーションが起きているのです。

しかし、データを有効に活用できるようになるまでには、いろいろと試行錯誤があります。まず重要なのはデータの集積です。レシピに関し、クックパッドと

いう名称のサイトがあります。サイトのスタートは1998年です。しかし、クックパッドが株式公開にこぎ着けたのは2007年です。同サイトのレシピ数の推移を見れば、公開まで時間がかかった理由が良く分かります。最初はなかなかレシピが集まらなかったのです。ところがレシピがある程度集まり、利用者がその価値に気付くとレシピの集積が加速しました。モバイルの発展がこれを後押ししました。十分な量が集まらないと、データの価値は高まりません。クックパッドの例は、データ集めの苦労を物語っています。

もう一つ大事なことがあります。それはイノベーションに挑戦することです。農業の例で述べたように、現在はデータ活用がイノベーションを促進する時代です。自分が世の中を大きく変える、という意識を持つことが重要です。今まで、日本のイノベーションの多くは、改良型のイノベーションでした。すなわち、燃費のいい車を作る、より速度の速い通信網を作るというように目標が最初から決まっており、それを上手に実現することが重要でした。

ところが、これからのイノベーションは違います。新しい価値の発見が重要なのです。アップルのiPhoneはパソコンを携帯端末に変え、モバイルの世界にパラダイムシフトを起こしました。このように、現在では今までにない価値を発見する型のイノベーションが増えています。iPhoneは短期間にパラダイムシフトを実現しましたが、データ活用によるイノベーションでは



※4 リモートセンシングデータ

人工衛星や航空機など地上から離れたところから、大気や地表を観測する技術によって得られた情報、天気予報にも活用されている。

人々の気付きが必要で、少し長い時間をかけてパラダイムシフトを実現する機会が多いように感じます。

変革の実現には、いろいろな切り口からビジネスを考えることが必要です。データ活用は、そのためのツールの一つなのです。ビジネス変革の仮説を作り、それを検証し、仮説が正しいのかどうかを確認しながら変革を進めていく。このプロセスを何回も繰り返すことが、イノベーションにつながるのです。そして、このプロセスの中で検証のためにデータを活用するのです。

イノベーションについて、もう一つ理解しておかなければならないことがあります。それは、発見型イノベーションの大半は失敗するということです。1,000に三つとか10に一つとか言われています。それが現実なのです。問題なのは、日本の企業や組織の多くが、失敗を恐れる風土を持っていることです。これを変えるには、「Nice Try!」という考え方を徹底することが必要です。挑戦する心をたたえ、失敗をとがめないマインドが重要です。

世の中には、失敗するのが嫌で何もしない人が大勢います。確かに今までは、チャレンジして失敗するよりも、何もせずに失敗しない方が出世には有利だったかもしれません。しかし、ビッグデータの時代には、これが変わります。収集したデータを分析することにより、失敗しても何が間違っていたかが分かり、次の挑戦で成功する確率が高くなるからです。新しいことに挑戦することで、次に進むべき方向が次第に見えてくるのです。逆に何もしないと、変化の方向が分からなくなり、対応が遅れる危険性があります。したがって、ビッグデータ時代の成功の反対は、失敗ではありません。何もしないことなのです。

皆さんにお願いしたいのは、マインドを変えビッグデータ活用に、そしてイノベーションにチャレンジすることです。もともと、地方にはイノベーションマインドに富んだ人が多いと感じます。新しいことに果敢に挑戦し、地域の活性化に結び付けて下さい。

## 講演 1

### LOD<sup>※5</sup>つながる “データシティ鯖江” ～オープンデータによるITのまちづくり～

鯖江市が目指すまちづくりには三つの柱があります。

一つ目は、「市民協働のまちづくり」です。今まで行政がやっていた公共の33事業を、今年は市民の皆さんに行ってもらおうことになりました。全部で700ぐらいの事業なのでまだまだなのですが、新しい取り組みです。



牧田 泰一 氏  
福井県鯖江市政策経営部情報統括監

二つ目は、「若者・学生とのまちづくり」です。私たちの市では、学生を対象に自分が市長になって鯖江市の活性化を考える「鯖江市地域活性化プランコンテスト」を毎年開催し、多くの学生が参加しています。

また、1月には「おとな版地域活性化プランコンテスト」を行いまして、その時の提案から「JK課プロジェクト」が立ち上がりました。これは今まで行政に一番遠いといわれていた女子高校生に、まちづくりに参加してもらおうという試みで、全国紙にも取り上げていただきました。

三つ目は、「ITのまちづくり」です。メガネ、繊維、漆器に続く4番目の産業にしたいという思いから、ITのまちづくりを進めています。最近ではオープンデータが脚光を浴びてきましたが、私たちは「ITで市民生活を豊かにする」という考え方でオープンデータなどIT推進に取り組んでいます。

オープンデータの定義ですが、行政情報は今まで、告示であったり、広報誌であったり、テレビであったり、いろいろなメディアにお知らせしたいことだけをお知らせしてきました。それで、欲しい情報は情報公開請求してくださいと、どちらかという受け身でしたが、Webの時代になってより積極的に市が所有するデータを公開しようという新たなデータの取り組みを始めました。オープンデータは、まず、機械が読め

※5 LOD (Linked Open Data)

世界中にくもの巣のように張り巡らされたWWWのリンクを持つデータの「公開」と「共有」をより効率的かつ実践的に行うための技術。

るデータ形式で公開する必要があります。もう一つは、「CC<sup>※6</sup>BY」、すなわちCC (Creative Commons) のBY (表示) だけをしてほしいというライセンスです。この二つが重要で、Web時代のデータ公開のあり方についてはこれが標準になりつつあります。

電子行政オープンデータ戦略の中では、意義、目的として三つ掲げられています。まずは、行政の透明性と信頼性の向上、国民参加、官民協働、鯖江市のような市でいうと、市民協働です。もう一つは、経済の活性化、行政の効率化です。稲田先生のお話の中では「イノベーション」です。例えば、行政の透明性は、今まで出していなかった情報を出すので、間違いなく透明性が確保されています。そして、市民参加です。市民参加は、まずは今まで活用されていなかったデータを出して、自由に使ってくださいということです。また、それから作られた分析やアプリは、多くの皆さんに使っていただきます。最終的には、市民の皆さんの利便性が上がるという考え方で、今まで行政に関与していなかった人でも、アプリという形で市民生活の向上に寄与していただくということです。ここまでの二つは鯖江市で実感していますが、経済の活性化や行政の効率化というところはまだこれからと思っています。

鯖江市のオープンデータの取り組みを担当する株式会社jig.jpの福野泰介さんから、Web時代の新しいインフラなので、行政からネットの中にどんどん整備していく「データシティ鯖江」の考え方を提案してもらいました。市長は、市民主役条例のこともありますし、若い人の意見を聞くことについてはすごく前向きな人で、「ぜひやってみたい。研究したい」ということでした。このとき市長が言ったのが、「いいことも悪いことも公開して、市民の皆さんに相談して、一緒に悩み、一緒に喜んでもらう」ということです。

情報データ媒体の進化といえ、情報を声で伝えるやり方、紙で伝えるやり方があります。そして電波や通信では、情報量が多く距離感が大きく即時性があり、しかも地球の裏側まで伝わるようになりました。これ

まで情報は独占的でしたがWeb時代になって、これらの項目に関してほとんど全部優位性が出てきています。例えばSNSで申し上げますと、距離感はほとんどないですし、四六時中会っているような感覚にさえなります。このことで国家の政治が変わってしまうような事態も起きています。それほどすごいツールになってきています。

情報量というところでは、鯖江市のホームページには情報があふれています。検索すれば出てきますが、どれが欲しい情報なのかなかなか見つけられません。つまり、何回も何回もクリックしないと欲しい情報が出てきません。ということで、情報を機械に処理させるためのオープンデータという考え方ができるかと思っています。

また、端末の進化があります。今は当たり前のように、昔でいうパソコンをポケットに入れて、持ち歩いています。それこそ、高度なGPSなどのセンサを持った端末を、だれでも持てるようになりました。こういう時代が来たからこそ、オープンデータという考え方が出てきたと思っています。

今から2年半前ですが、「トイレ情報」が一番身近なデータということで40件ぐらいのデータを公開しました。この情報からスマホで、現在地から一番近いトイレを案内するというアプリケーションができました。

「目指すオープンデータ」ということで、Webの発明者であるティム・バーナーズ=リーが五つ星オープンデータを定義しています。Linked Open Dataというものです。汎用性が低い一つ星のPDFから、三つ星でCSV、四つ星でRDF、さらに高度な五つ星の最高級のLinked-RDFという形式まで、星が多いほど機械が読み取りやすくなります。今、鯖江市は三つ星から始めていますが、近々四つ星に移行しようと思っています。

現在、「データシティ鯖江」という鯖江市のサイトで、統計情報、施設の情報、避難所、観光情報、文化財、鯖江検索情報、地図、市内つつじバスの位置情報など

※6 CC (Creative Commons)

インターネット時代のための新しい著作権ルールで、作者自らが「この条件を守れば、作品を自由に使用できる」と意思表示するためのライセンス。受け手は条件の範囲内で二次利用可能となる。

の情報を、テストで45種類公開しています。それに伴って、民間では90種類以上のアプリが公開されています。ほとんどがオープンデータの提案者である福野泰介さんのアプリです。

鯖江市は公共施設のWi-Fi化<sup>※7</sup>を3年前にやりました。これは、Wi-Fiを受信できるような環境を「見える化」したものです。スマホを片手に、JR鯖江駅から西山公園まで情報を取りながら歩いていきます。

このようなアプリに市民の皆様が親しんでもらおうと、まずはお金と時間がある高齢者に、iPadでのプログラミング教室を開きました。この中から本当に自分でアプリを作るという人も出てきて、「お年寄りだからパソコンが苦手」ということは全然ないと思いました。

平成26年度の取り組みとしては、総務省の実証実験の中で取り組んだポータルサイト<sup>※8</sup>を実際に公開したと思っています。データを公開すれば、アプリもセットで使えるということで、多くの自治体に使っていただきたいと思っています。

ITは今後もっと市民の皆さんに近づきつつあって、行政にもいろいろな形で関与してくることが予想されます。これまでは行政の中ではなかなか活用されませんでした。東日本大震災で活躍したように今後は先端技術を行政内部で使えるような形を作っていく動きになっていると思います。

## 講演 2

### 先進事例から学ぶビッグデータの活用法

#### ～最先端の分析技術によるイノベーションの創出～



中村 慎二 氏  
日本電気(株)ビッグデータ  
戦略本部本部長

ビッグデータとは何かというところをいろいろな切り口でお話ししたいと思います。

身近な例ですと、最近、フォークにセンサがついていて、何回上げ下げしたかカウントされる製品が出ています。これは何のためにやっているかという、速くフォークを上げ

下げしているということは、早食いしているということで、早食いを抑制するためのセンサ付きフォークです。こんなものにまでセンサが付き始めたわけです。

自動車に至っては、センサの固まりです。非常に多くのセンサが付いていて、それらのセンサ同士がデータをやり取りして、中のコンピューターで制御したり、そこに来たデータを片っ端からクラウド<sup>※9</sup>に送って、その車がどういうドライビングをしているかを自動車メーカーは全部取っているわけです。ものを言わぬものがものを言うようになってきています。いろいろなものがデータの発生源になってきているという視点があります。

それから、それを行政で活用する上で、これはオープンデータに近い話だと思いますが、埼玉県の取り組み例があります。これも自動車から上がってくるデータで、よく急ブレーキをかけられている場所があると、急ブレーキが多いのは危険な運転をする人が多い道路だろうと推測されるということです。実際に県の職員が見に行くと、そこは見通しが悪いのにスピードが出やすいことが分かりましたので、道路が狭く見えるように路側帯の線を引いたところ、急ブレーキの回数が減ったということです。

また、農業や金融、犯罪予防などさまざまな分野での取り組みが始まっています。今までは人間が経験と勘で「明日はこうなる」といっていたのが、膨大なデータを機械に食わせてやると、もっと精度よく、もっと速く法則を導き出して、予言してくれたり、コントロールしてくれるというふうになってきているのが、ビッグデータにより生まれてきている世界です。

今、世界人口は70億人です。2050年には90億人となると予測されています。ただし、均等に増えるのではなく、人口は都市に集中するので、都市の集中率が70%に高まります。すると世界経済規模で4倍ぐらいになります。インフラのエネルギーは1.8倍ぐらい必要になり、水も1.6倍ぐらい必要になるということです。今の人間生活を支えている社会インフラがまる

※7 Wi-Fi化  
無線でネットワークに接続する技術のこと。屋外でも高速にネットワークに接続し、便利にインターネットを活用可能となる。

※8 ポータルサイト  
ポータルサイトとは、検索エンジンやリンク集を核として利用者が必要とする機能を提供するサイトのこと。ここではオープンデータを簡単に利用できるように関連情報を集めた実験サイトを指す。

※9 クラウド  
ソフトウェアやハードウェアの利用権などをネットワーク越しにサービスとして利用者に提供するために運用されるサーバ群のこと。メールは代表的な身近なクラウド。

まるもう一つぐらい必要になる時代になっていきます。当然それだけのインフラを運用していくリソースはあるのか、エネルギーはあるのかということになるので、そうなると効率よくインフラを動かして、もっと安全・安心に暮らしていけるような「新しい社会インフラ」が重要になってくるとNECは考えています。

原発のオペレーションにビッグデータ分析を役立たたという事例では、この原発では圧力、振動、流量といったセンサが約3,500個あり、そのセンサが一つ当たり1秒間で100個のデータを送信します。この分析エンジンは何をやっているかという、3,500個のセンサ全部を総当たりで関係があるかどうかを見ていくのです。ですから、3,500×3,499組の組み合わせを全部モデル化しています。

従来は個別のセンサの値を上限と下限で見ている、その閾値を外れたら異常値、という判断をしていましたが、この分析エンジンでは個々の値は正常値であってもセンサ間の相関関係をモデル化し、「いつもと違う挙動」を早期に検知します。この技術を使うと、ベテランの人と試験設備を比べたときに、このエンジンのほうが10数時間早く予兆を発見できたということです。実際本稼働に向けてこれを導入していただいたということで、この技術をNECでは「インバリエント分析技術」と呼んでおり、石油や化学などのプラント、橋や道路といった社会インフラの故障予兆検知へと適用を広げ、安全・安心な社会づくりに貢献していきます。

二つ目の事例として、電力需要予測について紹介します。こちらは「異種混合学習技術」を使っています。営業日や気温、天候といった多種多様なデータから複数の予測式を自動で導出し専門家の仮説なしに高度の予測を可能とする技術です。これは他の業種でも活用されています。例えば、コンビニの商品の需要予測です。おにぎりが何個売れるか、シュークリームが何個売れるかという需要予測技術です。今はこの技術を使って時間帯毎の一個単位のおにぎりの需要数を算出

し、発注までしています。

まとめると、価値創造が、ビッグデータの最終目的だということです。いろいろな手順があると思いますが、まずはどういう価値があるかを定義して、本当にその価値が分析によって見つけられるか、価値が発見できるかを検証して、できると分かったら導入することになります。

ビッグデータの難しいところは、この評価するとうか、発見フェーズがあるというところで、従来のITシステムの導入とは違います。

実際にビッグデータの取り組みをされたお客さまに聞くと、9割の労力、9割の作業は分析する前の準備作業にかかるらしいのです。集まってくるデータが足りないとか、来るデータに不備があるとか、抜けがあるとか、この店とあの店が出てきたコードが違うとか、そうしたところを全部合わせていくという前処理に大体9割の工数がかかるとお考えください。

また、特に最近、プライバシーの保護や秘匿など、データの扱いがデリケートになっていて、「準個人データ」みたいなものも検討されているようです。

弊社としては、このほか画像認識サービスも行っていますし、ビッグデータ・オープンデータの活用等に関して全国各地の協議会や団体にも参加しています。

今後も、このビッグデータ・オープンデータ活用に積極的に取り組み、社会価値創造に貢献していくために、さまざまな技術開発を行い、また体制をつくっていますので、気になることがありましたら、お声をかけていただければと思います。