

■ ビッグデータ (TBSラジオ「日本全国8時です」2012. 8. 30)

ここしばらくIT(情報通信)社会で話題になっているビッグデータについて、御紹介したいと思います。

IT業界は世界でもっとも変化が目まぐるしい社会で、この20年ほどを振り返っても、80年代の「ニューメディア」から90年代の「マルチメディア」、そして「インターネット」へと発展し、21世紀になると、人間がモノや場所と情報交換をする「ユビキタス技術」が騒がれました。そして2005年頃から「クラウド」、最近では「ビッグデータ」という流れです。

ビッグデータは、その言葉のように膨大な情報という意味ですが「3V」という言葉によって特徴が説明できます。

第一は「ボリューム (Volume)」、すなわち大量の情報ということです。

昨年、世界全体で1年間に創造されたデジタル情報は1・8ゼタ・バイトになったと予測されています。

ゼタは0が21個連なる数字ですが、近いものに例えてみると、新聞の朝刊に含まれる記事の20億年分という気の遠くなるような情報量です。

そのうちツイッターだけでも1日に12テラバイトのつぶやきが発生しており、1年では新聞450万年分に相当します。

第二が「バラエティ (Variety)」で、文字や写真や動画や音声などの人間が創るデジタル情報だけではなく、スマートフォンがGPSの信号を受信して割り出した位置情報や、Suicaが改札口を通過した人間を確認した情報など、種々雑多な情報が創り出され蓄えられていることです。

第三は「ヴェロシティ (Velocity)」で、人間が頭をひねって時間をかけて創り出す情報だけではなく、コンビニエンスストアで買物をした瞬間に、POSによって場所、時刻、品種、値段、性別などの情報が送信されてしまうという高速時代になっているということです。

このような洪水に対応する2種類の技術革新が登場し、洪水の宝の流れに変わりはじめたのです。

第一は大量の情報を記憶させておくハードディスクや様々なセンサーの値段が急速に安くなってきたことです。

記憶装置のハードディスクは、10年前であれば数百万円した数テラバイトの装置が現在では数万円になり、10年で数百分の1に下がりました。

さらに現在ではクラウド時代になり、自分のコンピュータに大量の記憶装置を設置する必要はなく、グーグルなどにアクセスすれば、無料で相当の記憶容量を使用する

ことができます。

また、最近のスマートフォンやデジタルカメラにはGPS機能が組み込まれていますが、これも以前は1000万円もする大型装置でしたが、いまでは値段を無視できるほどのチップになっています。

第二は蓄積された大量の情報を分析することのできるソフトウェアが開発されたことです。

膨大な情報が蓄積されていますと言われても、新聞で何億年分というような情報は人手ではもちろんですが、スーパーコンピュータでも分析することは容易ではありません。

ところが、「ハドゥープ」を代表とする大量の情報を分散処理によって迅速に分析するソフトウェアが開発され、蓄積された情報が宝の山になりはじめたのです。

それではビッグデータで何が実現するかですが、多くの方が経験しておられるのは、アマゾンで本を検索すると「この商品をチェックした人はこんな商品もチェックしています」という文書が表示され、内容が似たような本や、同じ著者の別の本などが示される例です。

どうして、このようなサービスが可能かという、アマゾンを利用する人の検索したり購入したりした数百ペタバイトの情報がすべて蓄えられており、それらを分析した結果なのです。

グーグルでも検索のキーワードを間違っ「コロンベス」と入力すると「もしかして:コロンブス」という案内が出てきますが、これも多数の入力情報を分析した結果、多分、間違いだろうと推測しているからです。

この技術は、さらに将来予測にも利用されます。

アメリカには「フライトキャスター」というサービスがあり、予約しようとする航空便が1時間以上遅れる可能性は何%かを示してくれます。

これは過去の航空便の膨大な運行情報や気象情報を付き合わせて、どのようなときにどのくらい遅れたかを推測して結果を示しているのです。

このような分析をビジネスに応用すれば、顧客を抽出してアンケート調査をするよりも、過去の行動の履歴を大量に集めて分析した方が正確な需要が把握できますし、アマゾンの書物の推薦のように、一人一人の希望に対応することも可能になります。

そこで「データ・イズ・ニュー・オイル」という言葉が登場し、データこそが情報社会の原油で、それを「データ・マイニング」すなわち発掘し、分析すればビジネスの機会はいくらでもあるという時代になってきました。

アメリカでは連邦政府が200億円の予算を投入して「ビッグデータ研究開発イニシアティブ」という研究を今年の3月から開始し、情報社会の新しいビジネスを開発

しようとしています。

日本にも、いくつかの小規模な研究会がありますが、国家戦略はなく、遅れている情報産業がさらに遅れる心配があります。

もうひとつの心配は個人情報の分析によるプライバシーの侵害です。

スマートフォンで写真を撮影すると、自動的にGPS情報を記録する機能が内蔵されていますが、不注意に使用すると浮気がばれるという程度はまだしも、電子機器に接した行動のすべてが記録され分析される時代になると、まさにジョージ・オーウェルの「1984」の時代にもなりかねません。

この問題は、アメリカ議会では議論されはじめていますが、日本では話題にもなっていません。

フェイスブックやツイッターで自らプライバシーを積極的に公開している方々はともかく、知らない間にプライバシーを侵害されないように、最新の情報通信サービスの利用には知識と注意が必要な時代です。