

新ビッグデータ・システムの商品開発とその最適活用への挑戦

－IoT時代の事業新展開の検討－

㈱インフォコーパス 取締役 CTO 佐治信之氏

まず見ていただきたいのが、Twitter のログです。だれでも取れます。GPS 付きでツイートすると、実は誰がというのは正確には分からないが、〇〇さんがここで何を喋ったというデータが全部取れる。だれでも取れる。1 年間のデータをまとめてみました。2 時から 4 時、4 時から 6 時までの点の濃くなり具合、昼に向かって人が地方から集まってくる。12 時辺りは濃い。みんなが昼休みにツイートするから。そういった情報が実は誰でも取れる。一番点が少なくなるのは夜中の 3 時か 4 時で、都心の人間が夜遅くまで徹夜で仕事したり遊んだりしている。こういうことがわかってしまう時代です。これは 1 億ログくらいしか無く、いまのビッグデータの基準から言うとても小さい。今はテラ級の情報がある。今は一人のデータをつなげていくと、その人の行動が見えてしまう。そんなこと恐ろしくてできない、それをどうやって匿名化するかという話も重要になってきます。

そういったビッグデータの時代のデータ活用は実際どうなっているか。データの信頼性はセンサーで得られるデータはノイズも乗ってますし、場合によっては予期しないデータが紛れ込んでいるかもしれない。信頼性というものをどう担保していくかが重要である。一方、せっかく集めたデータをちゃんと保全していないとあとで活用できない。そういう意味で保全性が重要です。

それを貯めておいただけではダメで、データを分析してそこから価値を得なければならない。それを自社の中で閉じるのか、それともデータとして他の会社に売るとかという話になると、今度はそれを二次利用するということを前提にした加工、ここまで考えていく時代になってきた。

一方で、既に取れている情報がある、購買データ、POS レジの情報とか、在庫情報とか。それ以外に色々なデータが取れるようになってきた。一つは、ヘルスケア、生体情報、皆さんの脈拍や血圧などの健康管理のデータが、そこだけで閉じるのか、公共事業のデータとして使われるようになるのか、いろいろな面がある。先ほどの位置情報、環境情報、これはセンサーで得られた情報が多い。こういったすごく多様な情報が使える。では、と言ってそれを自由に放っておくとプライバシーを侵害することに繋がる。

そういったデータを扱っていく上で、データを上手く扱っていけばビッグデータになるが、データを集めるのは、いわゆる Internet of things です。そういった情報を集めてくる意味をもう一度考えていくべきかと思います。当然ながらデータを集めれば良いというわけではなく、IoT やビッグデータと言うのはあくまで判断です。重要なのは、何をやるのか、何を知りたいのか、という目的があって初めてデータを IoT で価値に変えていくというステップだと思う。一番顕著なのが故障検知です。温度と湿度と振動のデータを取っていて、実は湿度が本質的だったとわかってきた。経験則としてあるが、それをデータとしてわかってくる時代になってきた。逆にデータからインスパイアされているということですね。