

Computer Report

Vol. 53 No. 11 11月号 (通巻 710号)

はじめの言葉

■アプリケーション業務処理を中核にした情報システムの展開が一段落すると、集積／蓄積されたデータの分析／加工処理が注目される。業務系処理が定型的なものであるのに対して、情報系処理は非定型なものであり、情報の高度化利用などという表現もされてきた。対象となるデータも、定型処理（ルーチンワーク）用に対して、非定型処理（ストックワーク）用データが使われるということで、整理／区別されてきた。

■しかし昨今市場で喧伝されているビッグデータ分析のススメでは、こうした対象データを整理／区別することもなく、数多あるデータ群をフル対象にした分析／加工処理であるように感じられる。実に曖昧な点が多い。時の勢いというものがあるのは分かるが、インターネット上に存在する、限りなく無制限なデータを対象にしているような勢い（？）である。インターネット上のデータを片っ端からアクセスすることをススメているかのようだ。

■果たして、昨今のビッグデータ分析とは、どれくらいの範囲までのデータを対象にしたものかは判然としないが、定型／非定型を問わず、組織内部で発生／収集されたデータ以外に、外部から入手されるデータも含めたデータを対象にしているものようだ。その外部データの収集場所としてインターネットが想定されているとしたら、あまりにも安直すぎる気がする。ガベージイン／ガベージアウトを懸念する。

■言うまでもなく、ガベージイン／ガベージアウトとは、不正確なデータを入力すれば、不正確なデータが出力される。出力の質は入力データの質次第だという経験則のことである。そもそも、ルーチンワークはユーザー組織に所属する、ほぼ全員が関わる仕事であるのに対してデータ／情報分析作業は、ほんの限られたプロフェッショナルの専門領域の仕事である。一般論として、ルーチンワーク現場の担当者の仕事ではない。

■さらに、いたずらなインターネットでのデータ収集／情報アクセスは、企業等ユーザー組織にとって実に大きなリスクを伴う作業になる可能性を秘めていることを留意したい。インターネットは、誰にでも自由なデータ／情報アクセスを解放している一方、様々なユーザービヘイビアを監視している勢力が存在し、そういう勢力に、限りなく重要なデータ／情報を与えてしまっているという事実を認識しておく必要があるだろう。

■とにかく、ビッグデータの分析作業は、組織の不特定多数のユーザーが関与する類のものではない。また、組織人によるインターネット情報への「へたなアクセス行為」は、入手するデータ／情報以上に、内部情報を外部に漏えいさせる行為となってしまう可能性があることも想定しておきたい。便利さにかまける無意識な行為がもつでの、巧妙なフィッシングトリックなどにかからぬ用心をしたい。

■ビッグデータ分析にしてもそうだろうが、データ／情報分析は意思決定者支援のためという位置付けである。意思決定者は、経営トップに始まり、案件の内容次第では、どのような部署部門、あるいは地位／ポジションにも存在する。それだけに、意思決定者によるインターネットアクセス行為に関する情報は、外部の競合相手にとって、非常に重要かつ貴重なものである。それを察知する側とされる側との差異は大きい。 (藤見)