

## データリテラシー

ある TBS の番組で、法政大学の学長が、このところの文書改ざん問題に関連して「法政大学でもデータリテラシーの必要性を唱える教員が出てきた」と発言していました。2001 年に『データリテラシー』（共立出版）でデータリテラシーを提唱し、データリテラシーという総合科目を大学院生に講義してきた身としては心強い限りですが、「情報リテラシー」に近い意味で用いられているのかもしれませんが、「情報」よりも「データ」の方が少しは実体感がありますからなのかもしれませんが、『データリテラシー』にも書いたように、リテラシーは「読み書き能力」が転じて「教養」を意味するようになりました。データリテラシーは「データに関する教養」ということになります。昨今の事件を見るまでもなく、これからの世の中で「データに関する教養」がないと、さまざまな困難に巻き込まれるだろうことは皆さん容易に想像がつくことでしょう。しかし「データてなんだ」をはっきりさせておかないと、データリテラシーといっても単なる精神論、お題目にとどまってしまう。議論も噛み合いません。例えば「書籍」はデータでしょうか？ コンピュータサイエンスでは「コンピュータで扱えるものはなんでもデータと呼ぶ」ことにしていますので、確かに書籍もデータです。ある先生が「本には読んだ方がよい本、読まなくてもよい本、（読むと知力が低下するので）読むはいけない本の 3 種類がある」と喝破されましたが、確かに「書籍のリテラシー」もあるかもしれません。ほとんどの書籍のページ数が 16 か 8 の倍数になっていることもそうでしょう。しかし「書籍」は教養の源泉の一つですので「教養の教養」ということになってしまいます。これは半分冗談ですが。

『データ分析とデータサイエンス』では、『新明解』の「推論の根拠となる事実を記号で表現したもの」をデータの広義とし、「分析の対象となる変量の値の並び、さらにはその集まり」を狭義としています。残念ながら高等学校の必修科目で習うデータは未定義のままですが、明らかに後者です。このように枠組みを定めて初めてサイエンスとして成立します。

データサイエンスの枠組みでのデータリテラシーは、このようなデータを扱う上での教養ということになりますが、ポイントは「思い込みやこだわりからの自由」にあります。「1 から 10 を知る」はスマートな人の代名詞ですが、こだわりを生むことと隣り合わせです。一度こだわりを持つとそこから抜け出すのはそう容易ではありません。それを助けてくれるのがデータです。「データに裏切られる」は「たんなる思い込みだった」ということに他なりません。データの前では正直になれることを学ぶのがデータリテラシーです。スマートなあなた、データを「自分の考えを正当化する道具」と思っていないでしょうか？

2018-03-21 里程