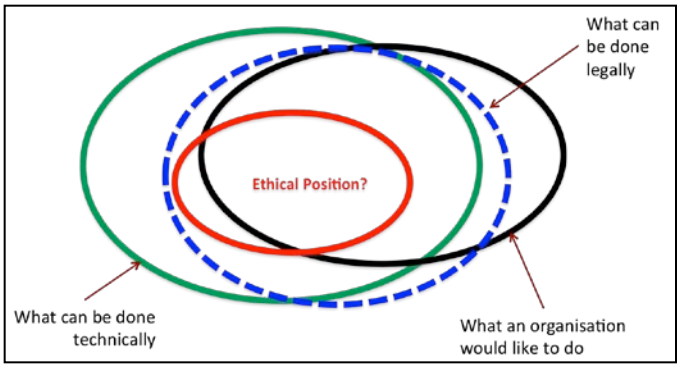
Ethics for big data and analytics

Mandy Chessell

Big data and analytics technology can reap huge benefits to both individuals and organizations – bringing personalized service, detection of fraud and abuse, efficient use of resources and prevention of failure or accident. So why are there questions being raised about the ethics of analytics, and its related technology, Big Data?

The technology itself is inherently ethics-agnostic. But it does push the art of the possible to new limits in terms of:

- The availability of a wide range of data from many sources.
- The ability to cheaply correlate this data to understand a bigger picture.
- The accuracy with which an individual can be identified and targeted.
- The ability to pinpoint someone's location for contextual insight and surveillance.
- The application of this new insight to a wide range of activities and actions.
- The operation of the insight in real-time or near real-time.



Laws and regulations guide organizations, particularly around privacy and the use of data, defining the current "no-go" areas for an organization. However, recent advancements in analytics and big data technology has widened the gap between what is possible and what is legally allowed, changing the balance of power between individuals and the data collectors. Within this gap are new opportunities alongside the risks of public relations disasters and unintended consequences. And it is within this gap where the ethical questions around what is acceptable are raised.

As an organization looks towards applying analytics and big data to enhance the way they operate, how do they know that their use of this technology is ethical?

At its core, an organization is "just people" and so are its customers and stakeholders. It will be individuals who choose what to organization does or does not do and individuals who will judge its appropriateness. As an individual, our perspective is formed from our experience and the opinions of those we respect. Not surprisingly, different people will have different opinions on what is appropriate use of big data and analytics technology particularly – so who decides which is "right"? Customers and stakeholders may have different opinions on to the organization about what is ethical.

This suggests that organizations should be thoughtful in their use of this technology; consulting widely and forming policies that record the decisions and conclusions they have come to. They will consider the wider implications of their activities including:

Context – For what purpose was the data originally surrendered? For what purpose is the data now being used? How far removed from the original context is its new use? Is this appropriate?

Consent & Choice – What are the choices given to an affected party? Do they know they are making a choice? Do they really understand what they are agreeing to? Do they really have an opportunity to decline? What alternatives are offered?

Reasonable – Is the depth and breadth of the data used and the relationships derived reasonable for the application it is used for?

Substantiated – Are the sources of data used appropriate, authoritative, complete and timely for the application?

Owned – Who owns the resulting insight? What are their responsibilities towards it in terms of its protection and the obligation to act?

Fair – How equitable are the results of the application to all parties? Is everyone properly compensated?

Considered – What are the consequences of the data collection and analysis?

Access – What access to data is given to the data subject?

Accountable – How are mistakes and unintended consequences detected and repaired? Can the interested parties check the results that affect them?

Together these facets are called the ethical awareness framework. This framework was developed by the UK and Ireland Technical Consultancy Group (TCG) to help people to develop ethical policies for their use of analytics and big data.

Examples of good and bad practices are emerging in the industry and in time they will guide regulation and legislation. The choices we make, as practitioners will ultimately determine the level of legislation imposed around the technology and our subsequent freedom to pioneer in this exciting emerging technical area.

About the Author: Mandy Chessell is an IBM Distinguished Engineer, Master Inventor and member of the Academy Leadership Team.