Taking advantage of Big Data revolution

DR HEZRI ADNAN

BITS AND BYTES: *Apart from training more data scientists and encouraging better sharing of information, there is a need to redefine what secrecy is, and is not, in the business of governing*

AMONG those running for the 2016 US presidency, Donald Trump has the biggest social media following with more than 14 million netizens. But, the pioneer status in harnessing the power of social media to rally for political support is not his.

In fact, it was Barack Obama's during the 2008 Presidential election. His reelection campaign in 2012 was even more data-driven. Relying on social media, his savvy political machine included a well-funded information technology team of 165 people. They worked nearly two years to collect data on who his target constituents were - the strong ones and fence-sitters.

Data from social media gave Obama the ability to hear what every individual and group was speaking and understand what the message was and what the positions were. From there, he was able to craft speeches to win the people over. This laser-like approach had ensured that he reached the right people with the right message to a larger group of voters.

A lot has changed since 2008. Back then, Facebook had a membership of only 5,000 people, while Twitter was in its infancy. Now, we are fully digital people. So much of what we do begins and ends as digital data. Not only do we upload photos to Facebook and Instagram, we send and receive email, surf the Internet for work or to pay bills, make online purchases and chat with friends.

All these information lead to large data sets stored in computing cloud, databases and warehouses. The increasing data flows, in social media and from sources such as text, web and sensors, can no longer be measured in the range of terabytes (10^{12}). Instead, they have grown to the scale of multiple petabytes (10^{15}), exabytes (10^{18}) or zettabytes (10^{21}).

Just to store one petabyte would take 223,000 DVDs with the storage capacity of 4.7 gigabytes each while five exabytes is equal to all the words ever spoken by mankind.

The increased volume, velocity and' variety of data has led observers to proclaim a new era of Big Data. It parallels some of the most significant milestones in the history of computing, equivalent to the invention of personal computers in the 1970s, the .World Wide Web in the 1990s, and social media in the 2000s. For these reasons, the World Economic Forum dubbed data a new "asset class" like currency or gold while observers have mused that data are "the new oil".

But, Big Data is useful only if we use the data in analyses. The grand challenge is going to be the ability of humans to make sense of the huge data streams using the Big Data Analytics (BDA) technology.
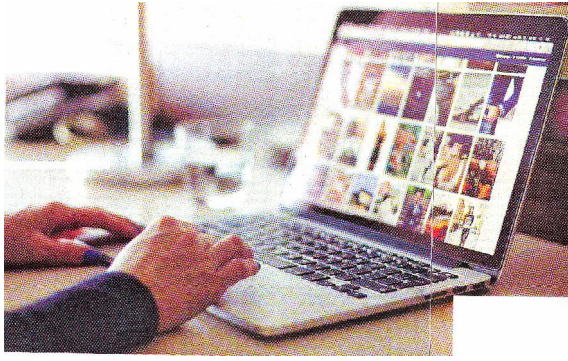
We are fully digital people. So much of what we do begins and ends as digital data. Not only do we upload photos to Facebook and Instagram, we send and receive email, surf the Internet for work or pay bills, make online purchases, and chat with friends.

What makes BDA or analytics particularly powerful is its capability of identifying patterns that are not detectable by the human mind. With appropriate modelling algorithms, BDA allows practitioners, analysts and researchers to also predict the onset of trends far earlier than was previously possible.

The analytics revolution is already generating value from the waves of new data for business, government and machine operation. The business model of Uber, the smartphone-app based taxi service, is firmly rooted in BDA. It is based in the Big Data principle of crowd sourcing and uses real-time analytics to predict when the driver will arrive to pick up a passenger and how to charge them.

The last time the American spy service CIA added a directorate to its agency was in 1963. Last year, it set up its news Directorate for Digital Innovation to tap into the power of analytics. The aim is to turn the data the agency collects and monitors into useful insights for its secret agents and the nation.

Teradata, a global leader in BDA solutions, provides Siemens the cutting edge data-driven analytics. This ensures its transportation around the globe is fast, reliable and energy efficient.

All sensor data from high-speed trains go into Siemens Teradata Unified Data Architecture whose analytics can predict certain failures way in advance, avoiding service interruptions for its customers across the world.

Malaysia is also encouraging a number of government agencies to use its Public Sector Big Data framework. Recent pilot applications of BDA include the prevention of car thefts by the Royal Malaysia Police and sentiment analysis on citizen patriotism by the Malaysian Administrative Modernisation and Management Planning Unit.

The excitement of Big Data, however, comes with many misconceptions.

FIRST, reporting or statistic generation is not the same with a meaningful analysis. Too many organisations equate business intelligence reporting with advanced analytics, often leading to wrong diagnoses and decisions.

An insightful end-to-end analytics team will combine computing prowess with subject matter expertise. Such an approach enables the team to interpret data trends and infer potential actions and steps that can be taken to address a problem.

SECOND, a traditional information technology policy will not drive a transformative analytics culture. Instead, BDA should be considered a pillar of the Malaysian public policy, and not an optional add-on.

To leverage Big Data fully, we have to focus on people, not tools. This means not only training more data scientists and encouraging better information sharing between agencies, we must also redefine what secrecy is, and is not, in the business of governing.

The writer is director of Technology, Innovation, Environment and Sustainability at Institute of Strategic and International Studies (ISIS) Malaysia