BIG DATA AND ITS OPPORTUNITIES FOR BUSINESSES

大数据及其为企业带来的机遇

By

Shim Zhou Hong
Sum Kum Mooi

Socio-Economic Research Centre (SERC)

July 2016

This research paper was commissioned and funded by ACCCIM SERC Trust and managed by the Socio-Economic Research Centre (operating under SERC Sdn. Bhd.).

Mr Shim Zhou Hong (zhshim@acccimserc.com) is a Researcher of SERC; Ms Sum Kum Mooi (kmsum@acccimserc.com) is the Senior Researcher.

Tel: 603-42603116 Fax: 603-42603118

Contents

List of Tables

List of Figures

Executive Summary

We live in a Big Data environment today. The term “Big Data” refers to massive amounts of data in multiple formats generated inside and outside organizations on day-to-day basis. The three main characteristics of Big Data are commonly referred to as 3V's: (1) volume, i.e. massive amounts of data; (2) variety, i.e. structured, semi-structured and unstructured data generated from multiple sources; and (3) velocity, i.e. data generated at fast speeds. The explosive growth of semi-structured and unstructured data — just think of social media postings — is a key factor that lies behind the Big Data phenomenon. It is estimated that semi-structured and unstructured data account for 80% to 90% of all data. Traditional data technology cannot handle these data efficiently and effectively, and this implies that organizations were mainly relying on just 10% to 20% of all data to guide their decision making in the past. The emergence of Big Data technologies has enabled organizations to leverage Big Data and discover previously hidden insights.

Big Data can provide richer, more granular and faster insights to power smarter and faster decision making in a wide range of business problems. Procter and Gamble (P&G) searches for new product ideas from customers' comments on social media; Merck, a US-based pharmaceutical firm, leverages Big Data and predictive analytics for deeper root-cause analysis to improve production yield; United Parcel Service (UPS) analyzes real time road traffic data for route optimization; and Celcom delivers targeted promotions to customers based on their actual behaviors. Several reputed research centers found that firms with higher business analytics competency outperform those with less. International Data Corporation (IDC) found that analytically oriented firms are 20% more likely to be among the most competitive within their respective industries. According to Gartner, firms that use predictive analytics may increase profitability by 20% by 2017. A joint research by MIT Center for Digital Business and McKinsey's Business Technology found that data-driven firms perform on average 5% more productive and 6% more profitable than their competitors. These findings show that there is a strong case for Big Data adoption.

Big firms are generally good at leveraging data in their decision making. In contrast, small- and medium-sized enterprises (SMEs) are generally less data-driven and rely more on gut feeling and past experience in their decision making, hindered by smaller data sets and budget for technology investments. Nevertheless, in today's Big Data environment with increasing availability of quick-to-deploy, affordable and flexible technology products, SMEs have better and cheaper access to data, and this represents a big opportunity for them to develop Big Data capabilities to win in the ever more competitive marketplace. SMEs can leverage cloud-based products which require no hardware, no setup time and can be

deployed with just a few mouse clicks. These products also have affordable and flexible pricing plans that allow users to make incremental increase as their needs grow. Some cloud-based analytics products also provide access to external data, e.g. Google Analytics provides access to web traffic data and IBM Watson Analytics provides access to Twitter.

Big Data adoption is picking up across the world with more and more firms jumping on the bandwagon. IDC forecasts the global Big Data technology and services market to grow robustly at a compound annual growth rate (CAGR) of 23.1% from 2014 to reach USD48.6 billion in 2019. However, there is a huge Big Data talent shortage. The US alone is expected to face a shortage of 140 000 to 190 000 workers with deep analytical skills as well as 1.5 million managers and analysts with know-how of Big Data analysis to make effective decisions by 2018. As a late comer in the industry, Malaysia too lacks Big Data talent. According to Malaysia's Minister of Higher Education, there were about 4 000 Big Data scientists in 2014 and an additional 12 000 will be needed in the next five years. SMEs are generally seen as less attractive employers compared to big firms. It will therefore be less attractive for SMEs to embark on Big Data projects since it is useless to invest in the necessary infrastructure if no one is there to make sense of data and make insights actionable. SMEs which lack internal Big Data talent can consider engaging external consultants instead of developing an in-house team, or turn to Kaggle for help — an online platform that matches data-related requests with data scientists from around the world at a far more affordable cost.

It is a tough task to deliver value from Big Data, even for big, data-driven firms. Many big firms which are committed in Big Data and advanced analytics failed to achieve the big impact they expected. Nevertheless, there are also plenty of successful examples where firms have been able to improve their top-line and bottom-line performances. To increase chances of success, firms should develop a data-driven culture throughout their organizations; carefully evaluate the business case of their proposed Big Data projects; and clearly define the specific outcomes, i.e. think of what to ask of the data, how the firms will react to the answers, and what are the actionable operational measures.

Data is the new oil. Firms that are not data-driven risk of being outcompeted by those that are. The complexity of today's business environment warrants them not to only rely on gut feeling and past experience but to adopt a more data-driven approach in their decision making. Big Data can be the new strategy to help firms stay competitive and so they should invest in Big Data technology, talent and culture to prepare themselves for the Big Data era.

摘要

我们生活在一个“大数据”时代。简单来说，大数据是指巨量、多样的数据以快速度产生。大数据有三个主要特征，通常被称为“3V”：（一）容量“volume”，即巨量的数据；（二）多样性“variety”，即多源的结构化、半结构化、和非结构化数据；及（三）速度“velocity”，即数据产生的速度非常快。半结构化和非结构化数据，例如社交媒体数据爆炸性增长是促使大数据现象产生的关键因素。根据估计，所有数据中有高达 80%至 90%为半结构化和非结构化数据。可是传统数据科技并不能有效地处理这些数据。这意味着企业们过去在做决定时仅使用所有数据的 10%至 20%。大数据技术的出现让企业们能够充分利用他们所拥有的数据并发掘先前隐藏的真知灼见。

大数据分析可提供更丰富、更深入、及更快速的见解，以让企业决策者们能够做更智能、更快的决定。Procter & Gamble（P&G）通过分析顾客在社交媒体上的评论来寻找新产品创意；Merck，一家美国制药企业，利用大数据和预测分析来进行更深层的根本原因分析以提高生产效率；United Parcel Service（UPS）通过分析实时交通数据来优化路径；及 Celcom 分析顾客的实际行为以提供针对性服务。多家知名研究中心发现，拥有较高业务分析能力的企业的表现优于其竞争者。International Data Corporation（IDC）的研究发现，分析导向企业成为行业佼佼者的机率比其竞争者高 20%。Gartner 估计到 2017 年使用预测分析的企业可提高盈利 20%。麻省理工学院数字业务中心联合 McKinsey 咨询公司的研究发现，数据驱动型企业比其竞争者的生产力平均高 5%、盈利平均高 6%。这些研究结果是支持企业采用大数据强有力的商业理由。

普遍上，大企业做决定时比较善用数据。相反的，中小型企业做决定时比较不基于数据，而是凭直觉和经验，这主要是因为他们拥有较小的数据集并限制于对科技投资预算。然而，在当今的大数据环境里，市场上有越来越多可快速投入使用、经济实惠、及更具弹性的云端科技产品，提供中小型企业更便宜、更实惠的方式来获取和使用数据。这对中小型企业来说是一个发展大数据能力的好机会，以便能够在这竞争日益激烈的市场中取胜。云端产品无需硬件设施、无需设置时间、并只需在滑鼠上按几下就可以开始使用，而且也具有实惠、灵活的定价，允许用户依据他们的实际需求调整用量。此外，一些云端产品也提供第三者数据，例如谷歌分析提供网络交通数据及 IBM 提供推特数据。

世界各地越来越多企业开始采用大数据。IDC 估计全球大数据科技和服务市场在 2014 年至 2019 年期间平均每年将增长 23.1%并达至 486 亿美元。可是大数据人才短缺。根据估计，截止 2018 年，美国将需要 140，000 至 190，000 名拥有高分析能力人才，及 1.05 百万名拥有大数据分析知识的主管和分析员。作为大数据的后来者，我国亦面对大数据人才短缺。根据我国高

等教育部部长，我国在 2014 年拥有 4，000 名大数据科学家，及将在未来的五年内需要额外 12，000 名相关人才。相较大企业，中小型企业一般上被视为不那么有吸引力的雇主。所以对中小型企业而言，投资大数据项目可能不是那么有吸引力，因为如果没有相关人才来分析数据和善用所得见解，投资大数据科技只会浪费资源。面对人才短缺，建议缺乏内部人才的中小型企业，与其培养本身内部团队，可以考虑雇用外部顾问，或通过互联网寻求相关人才，例如通过 Kaggle，一个互联网数据平台，来以更实惠的价格寻求全球各地数据学家的服务。

要从大数据创造价值并不容易。对大型、数据驱动型企业亦是如此。很多致力于大数据及高端分析的大型企业，并未能如预期般达到所谓的大价值。然而，也有很多成功利用大数据来提升企业表现的例子。为了提高使用大数据成功几率，企业因致力于培养内部数据驱动文化；仔细地评估所提议大数据项目的商业价值；及明确界定具体成果，即思考要对数据提出的问题、得到答案后应如何对应、及有什么可行的操作措施。

数据是新时代黑金。非数据驱动型企业可能被数据驱动型竞争者淘汰。在当今复杂的商业环境，企业们在做决定是不能再仅靠直觉和经验，而需善用数据。大数据可作为企业的新策略，以帮助保持竞争力。所以企业应考虑投资于大数据技术、人才、及企业文化，以作好准备面对这大数据时代带来的机遇和挑战。

1.0 Introduction

Big Data has risen to prominence in recent years as a result of decreasing cost of data infrastructure, new data processing and analytic technology, and increased availability of data talent (BCG, 2013). International Data Corporation (IDC) expects strong growth in Big Data adoption and forecasts the global Big Data technology and services market to grow at a compound annual growth rate (CAGR) of 23.1% from 2014 to reach USD48.6 billion in 2019 (IDC, 2015).

Countries worldwide such as Australia, Singapore, the UK and the US have been increasingly investing in Big Data to improve their public services and to encourage Big Data adoption in their private sectors. On the domestic front, our Government introduced a Big Data framework in 2014 with the objectives to increase Big Data adoption in both the public and private sectors and to make Malaysia a Big Data analytics hub in ASEAN by 2020 (Digital News Asia, 2014; Hanchard & Ramdas, 2014; The Sun Daily, 2015). A number of high-impact Big Data projects have been started, such as on dengue prediction and prevention, fighting organized crime and drug trafficking, and tax fraud detection (Malay Mail, 2015).

Firms across industries have also been increasingly investing in Big Data to drive their top-line and bottom-line performances. Several well-known research centers found that firms with higher business analytics competency outperform those with less. IDC's research found that analytically oriented firms are 20% more likely to be among the most competitive within their respective industries (IDC, 2012). According to Gartner, firms that use predictive analytics could increase profitability by 20% by 2017 (Hanmark, 2015). A joint research by MIT Center for Digital Business and McKinsey's Business Technology found that data-driven firms perform on average 5% more productive and 6% more profitable than their competitors (McAfee & Brynjolfsson, 2012). These findings show that there is a strong business case for Big Data adoption. On the domestic front, our private sector is lagging behind in Big Data adoption, due to lack of awareness and a wait-and-see approach as Big Data is still relatively new in the country (The Star Online, 2014).

Big firms are generally good at leveraging data in their decision making. In contrast, small- and medium-size enterprises (SMEs) are less data-driven and tend to rely more on gut feeling and experience in their decision making, hindered by budget for technology investment. Nevertheless, in today's Big Data environment with increasing availability of quick-to-deploy, affordable and flexible technology solutions, SMEs have better and cheaper access to data, and this represents a big opportunity for them to develop Big Data capabilities.

Against this background, this paper provides an overview of Big Data and outline some challenges as well as recommendations for SMEs. The remaining of the paper is organized as follows. Section 2 explains the main characteristics of Big Data, followed by an overview on different types of Big Data analytics in section 3. Section 4 provides some Big Data use cases and real world applications in product development, production, supply chain management, and marketing respectively. Section 5 discusses some challenges and recommendations for SMEs, followed by a conclusion in section 6.

2.0 What is Big Data?

The term "Big Data" can be confusing. It is new; everybody talks about it, lauds its enormous potential benefits, but raises fear of losing competitiveness if organizations do not adopt it, as digital is widely seen as the path for future growth; and it is wrapped in fancy, elusive terms such as Hadoop, MapReduce, machine learning, natural language processing and in-memory computing — mainly names of software and techniques for data processing and analysis.

In a straightforward sense, Big Data refers to massive amounts of data in multiple formats generated inside and outside organizations on day-to-day basis. The main characteristics of Big Data can be summarized as 3V's: volume, variety and velocity (see Figure 1).

Figure 1: Big Data's three main characteristics

Volume	Variety	Velocity
• Massive amounts of data are collected and stored by organizations • IBM says we collectively generate 2.5 EB (or 2.5 billion GB) of data daily	• Data in multiple formats from multiple sources • E.g. organizations collect internal transaction data (structured) and social media data (mainly unstructured)	• Data are generated at fast speeds • E.g. constant streams of credit card transactions and social media postings

Sources: Gartner; IBM

2.1 Volume

Big Data refers to massive amounts of data. While most people are familiar with the data storage units GB and TB (which are common storage limits of our laptops and external hard disks), the total amount of data we collectively generate daily today is in the order of EB (which is billion times larger than GB) (see Figure 2). IBM says on its website that about 2.5 EB of data are generated every day. The former CEO of Google Eric Schmidt once said that

5 EB of data generated every two days was equivalent to the same amount of data accumulated from the dawn of civilization up until 2003 (Siegler, 2010). It is not feasible to prove this statement, but it gives an idea of how massive data volume has become today. According to the EMC Digital Universe Study in 2014, the size of our digital universe (i.e. all the data we create and copy) in 2013 was 4.4 ZB (or 4.4 trillion GB) and is expected to double in size every two years to reach 44 ZB (or 44 trillion GB) by 2020 (IDC, 2014).

Figure 2: Ascending order of data storage units

Byte (B)	Kilobyte (KB)	Megabyte (MB)	Gigabyte (GB)	Terabyte (TB)	Petabyte (PB)	Exabyte (EB)	Zettabyte (ZB)	Yottabyte (YB)
= 8 bits	= 1 000 (10^3) bytes	= 1 000 000 (10^6) bytes	= 1 000 000 000 (10^9) bytes	= 1 000 000 000 000 (10^{12}) bytes	= 1 000 000 000 000 000 (10^{15}) bytes	= 1 000 000 000 000 000 000 (10^{18}) bytes	= 1 000 000 000 000 000 000 000 (10^{21}) bytes	= 1 000 000 000 000 000 000 000 000 (10^{24}) bytes

Source: The Oxford Math Center

Due to its sheer volume, Big Data requires a lot of data capacity and capability. For example, Walmart has accumulated historical data in the order of PB and collects some TB of new data every day; leveraging this massive amounts of data entails massive investments in Big Data technology, which may not be affordable for most organizations (SAS, nd).

The question is, how many organizations are actually using data sets that are large enough to require Big Data technology? According to an annual survey in 2015 (started since 2006) conducted by KDnuggets (a website for Big Data), most data users actually work with data sets that can be analyzed using simple analytical tools; majority of them work with data sets in the 10 MB to 10 TB range; the median data set is 30 GB; and merely 5% of them work with data sets in the PB range (Pafka, 2015).

2.2 Variety

Big Data refers to data in multiple formats, i.e. structured, semi-structured and unstructured, generated inside and outside organizations (see Figure 3). The rapid proliferation of semi-structured and unstructured data is a key factor that lies behind today's rising Big Data phenomenon. These data are said to account for 80% to 90% of all data (Dasgupta, 2015; Quantum, 2015; Spire, 2016). Traditional data technology can handle structured data efficiently and effectively but not semi-structured and unstructured data. This implies that organizations were mainly relying on only 10% to 20% of all data in their decision making in the past. With Big Data technology, semi-structured and unstructured can be harnessed more efficiently and effectively for richer and more granular insights.

Figure 3: Definitions and examples of different types of data

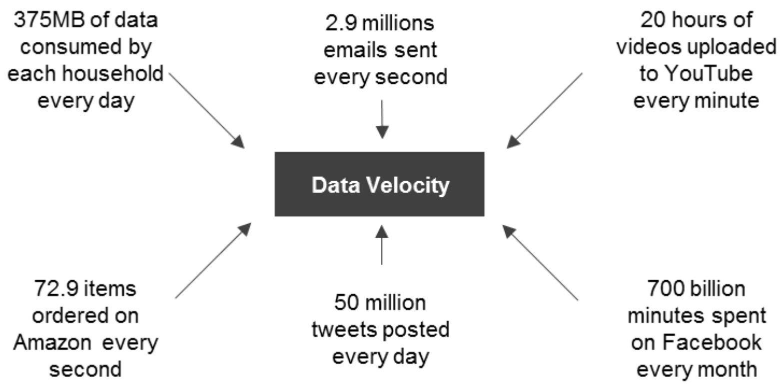
	Structured	Unstructured	Semi-structured
Definition	Data that can be grouped into a relational table, i.e. rows and columns within a relational database	Data with no specific structure and are in formats that cannot easily be indexed into a relational table	Data that, unlike unstructured data, contain some implicit structures but not enough as in structured data
Example	• Customer Relationship Management (CRM) system • Payroll data • Inventory data	• Emails • Video • Website content • Health records • Sensor data • Satellite images	• Web logs, e.g. IP address, request time, page requested • Social-media feeds, e.g. hashtag • Resumes, e.g. education, work experience

Sources: Loshin (2005); Anantharamn (2013)

2.3 Velocity

Big Data refers to data generated at fast speeds (see Figure 4), so technologies that are capable of capturing, storing, processing and analyzing them in an efficient and effective manner are needed. With traditional data technology, data move through systems in batch processes at pre-determined time intervals, so there are time delays between input and output; whereas Big Data technology is capable of handling high-velocity data for faster insights to power faster decision making (EITO, 2013).

Figure 4: Big Data's high velocity – Some examples



Source: IBM (2013)

Many business problems can benefit from real time Big Data application. Nongfu Spring, the largest bottled water producer in China, leverages Big Data technology to process its large volume of real time supply chain data and accelerate its 150 reports' generation time by 200 to 300 times faster; for example, it used to take 24 hours to produce a freight account settlement report but now just 37 seconds (SAP AG, 2012). Other potential applications include car manufacturers analyze streaming data from in-car sensors for predictive maintenance; payment networks analyze streaming transaction data for fraud; and e-retailers monitor real time website traffic to improve online marketing.

3.0 Big Data Analytics

Big Data is still data. Its uses are essentially the same as how data has always been used for, such as to identify consumer purchase motivation or bottleneck in production. However, its sheer volume and complexity exceed the capacity and capability of traditional data technology. Therefore, from a technology perspective, Big Data technology can be seen as a set of new, advanced technology solutions that are capable of handling Big Data efficiently and effectively.[1]

[1] Due to limitations of traditional data technology, organizations with Big Data can only leverage a smaller sample of it and mainly perform analysis on structured data using a small set of variables, and may further support it with more manual analysis of semi-structured and unstructured data. A traditional approach to analysis of semi-structured and unstructured data could take the form of reading all written communications (e.g. emails, tweets and reviews), watching all videos, and listening to all audio recordings, then manually converting emotional expressions and interpretations of the conversations into structured data, and then feeding the converted structured data into traditional analytic tools (Inside BIGDATA, 2015). The process is tedious, costly and error- and bias-prone. In contrast, with Big Data technology, the organizations can use all or a larger sample of its Big Data and integrate different types of data for analysis, e.g. they may integrate their internal operation data from Enterprise Resourcing Planning (ERP) system with semi-structured data from web-log files that identify customers' online behaviour and unstructured data from customer feedbacks in complaint emails or on social media.

Big Data technology can be divided into four layers: (1) infrastructure; (2) data organization and management software; (3) analytics and discovery software; and (4) decision support and automation software (IDA, 2012). This section will briefly describe four common types of Big Data analytics as follow:

- Data visualization - Understand what happened and is happening
- Predictive analytics - Predict what is going to happen
- Text analytics - Process textual data
- Graph analytics - Analyze connections between data

These analytics are not new, but as mentioned Big Data analytics have greater capacity and capability to process and analyze Big Data. Understanding what they are capable of doing allows us to understand better what kinds of insights and benefits may be gained from Big Data. Table 1 lists some providers of Big Data technology in Malaysia.

Table 1: Big Data open-source tools and vendors and training providers in Malaysia

Infrastructure	Analytics	Training	Open source
• Cloudera • IBM • SAP • Microsoft Azure • Amazon Web Services	• Google Analytics • IBM Watson Analytics • Tableau • Qlik • SAS Cloud Analytics	• Iconix Consulting • Abyres • Info Trek • iGen Technology • MaGIC Academy • MSC Malaysia MyProcert (SRI) – Data Science Massive Open Online Courses	• R statistical computing • Python programming • Weka machine learning algorithms • Apache Hadoop distributed data processing software • Apache Storm distributed real-time computation system

Note: For a more comprehensive list of players in Malaysia, access www.dropbox.com/s/0ow35i5vymuvf2m/Malaysia%20Big%20Data%20Landscape.pdf?dl=0

3.1 Data visualization

Data visualization enables users to uncover hidden patterns such as correlation and trend. Compared to traditional data visualization tools, Big Data visualization tools are capable of processing and visualizing huge and complex data sets at faster speeds for better and faster decision making.

A study found that managers who make use of data visualization tools are 8% more likely than their peers to access right information at the right time, enabling them to impact decision making (IDA, 2012). MIMOS, Malaysia's national research and development center in information and communication technology, used Big Data visualization tools to track how

GST sentiment correlated with the price of chickens and how routes travelled by police vehicles in Klang Valley correlated with crime rates in specific areas (Nadarajah, 2015).

3.2 Predictive analytics

Predictive analytics involves using mathematical and statistical techniques to exploit patterns found in data to predict the probability of future events occurring. Many organizations have collected multiple time series data stretching back many years in time. Big Data predictive analysis enables them to exploit the full value of their large data sets for forward-looking insights to drive smarter decisions. It has been used for a variety of business purposes, such as recommendation engine, sales and revenue forecasting, supply chain optimization, and modelling and simulation in product development.

A famous real world application is the Google's self-driving car project. The car is equipped with sensors to detect objects such as pedestrians, cyclists and vehicles as far as two football fields away in all directions. When on road, Google's software process both information from Google Maps and the sensors, predict what all the objects around the car do next, and then choose a safe trajectory for the car.

A more relevant example to businesses, particularly SMEs, is the use of predictive analytics for cost modelling by L&L, a medium-sized custom chemical compounding firm in the US. Prior to using predictive analytics, L&L built its cost models by generating charts for its over 40 input factors, visually comparing them, digesting the information, and then combining with past experience and gut feeling to estimate a composite number for the ultimate cost drivers. The work was tedious and subjective. It then leveraged open-source tools and cloud-based solutions for predictive analytics to develop, test and validate cost models. It managed to reduce the over 40 variables to a core set of four variables, enabling rapid estimation of the final cost drivers. This has not only reduced the cost forecast development time to just a few hours but also reduced subjectivity in decision making. (SimaFore, 2011)

3.3 Text analytics

In text analytics, linguistic, statistical and machine learning techniques are used to capture patterns in textual data so that meaningful information can be identified and extracted (IDA, 2012). Textual data are seldom used due to technology limitations but they may contain valuable information that do not exist in structured data. Therefore, it may unlock previously hidden insights that enable organizations to make better informed decisions.

It is particularly useful in analyzing customer sentiment and experience. For example, hotels usually receive large number of customer emails and their customer conversations are also increasingly taking place on the Internet where customers share and rate their experiences. They can use text analytics to crunch customer emails and online reviews to better understand the likes and dislikes of their customers. This can provide them better insights as to, for example, how to improve their customer relationship management or prioritize their investment spending. (Inside BIGDATA, 2015)

3.4 Graph analytics

Graph analytics maps connections between data in graphical format consisting of nodes and links, and provides insights into the strength of the connections. Compared to other analytics, it requires less coding and has faster and better performance, thereby reducing time and cost (Neo Technology, 2014). It can be used for single-path analysis, i.e. to understand causal process underlying an observed relationship; for optimal-path analysis, i.e. to identify optimized path; and for centrality analysis, i.e. to identify items/events that lie at the root of other surrounding patterns/events (IDA, 2012). Effective graph analytics needs to be able to handle large volume of, extremely varied, and often real time data (Hoskins, 2015). Thus, it has only been gaining importance with the breakthrough of Big Data technology that enables it to deliver the scale and speed that meet real world business requirements. Real world business use cases include network impact analysis, recommendations, sentiment analysis, fraud detection, geographic routing and logistics, and risk management (Wagner, 2014).

It has been used to identify fraudulent transactions and money laundering. The International Consortium of Investigative Journalists (ICIJ), a global network of investigative journalists, had obtained a massive and complex data set with details related to bank accounts of over 100 000 clients located in over 200 countries. It would be very time consuming and error-prone if the journalists had to analyze the data set using excel and made manual searches on Internet to try to identify connections between people and entities. Instead, ICIJ leveraged graph analytics to turn the data into graphical format consisting of over 275 000 nodes and 400 000 relationships, and then fed them into a web-based data visualization tool that provided convenient access to the journalists. This enabled them to analyze the huge data set more easily and faster, which would otherwise take a much longer time if they had to work manually and in silo. (Kepes, 2015)

It has been used to optimize logistics. United Parcel Service (UPS) leverages advanced mathematical models and a powerful system to process real time Big Data from in-vehicle sensors and other sources to provide real time route optimization updates to its drivers, and reportedly saved more than USD30 million a year in fuel cost (Ross, Beath & Quaadgras, 2013; Van Rijmenam, 2014). Similarly, DHL has developed a dynamic routing system that optimizes delivery routes in real time based on drivers' current orders and real time traffic data (DHL, 2013).

It has also been used to optimize cross- and upselling to customers. Walmart leverages graph technology to connect its massive amounts of complex buyer and product data to gain insights into its online shoppers' behaviors, preferences, and social networks for real time targeted product recommendations (Neo Technology, nd).

4.0 Big Data in business

Typical business objectives are to increase output, increase quality, reduce cost, increase sales etc. Firms have adopted various strategies — such as automation, lean management and outsourcing — to achieve these business objectives, but they are always in search for new strategies that can help them stay competitive in the ever more competitive marketplace. McKinsey has touted Big Data as the next frontier for innovation, competition and productivity (McKinsey, 2011a). A typical business value chain involves product development, production, supply chain management, and marketing. Each of the functional areas on the value chain has its own challenges. This section provides some examples of Big Data use cases as well as real world applications in each of these functional areas.

4.1 Product development

Product development entails high risk and cost. It is time consuming, ties up resources, and at risk for premature termination, and further, the final product may not be well received by the market. A data-driven approach to product development can reduce risk and increase chances of success. In today's Big Data environment, firms have access to many Big Data, such as sales, customer emails and social media. These Big Data can be mined for insights to guide better decision making in product development. Table 2 provides some Big Data use cases in product development.

A frequently cited real world Big Data application in product development is by Procter and Gamble (P&G), which is touted as having successfully leveraging Big Data in its product development. For example, it implemented the "Consumer Pulse" initiative which leverages advanced analytics to scan through comments on social media made by various stakeholders, categorize them by brand, and then feed them back to relevant managers, enabling quick reaction to what is going on in the marketplace. Further, by leveraging its customer insights on the deficits of the then dry cleaning industry and consumer household cleaning habits in the US, it introduced Tide Dry Cleaners, a dry cleaning franchise providing services that align with consumer preferences such as 24-hour pickup, drive-through service, and environmentally safe cleaning processes. (McKinsey, 2011b; Anastasia, 2015)

Table 2: Big Data use cases in product development

Use case	Description
Identify new product ideas that meet untapped customer demand	• Customer conversations on social media, complaint emails, call-center recordings etc. often contain constructive criticisms that are valuable sources of new product ideas. • P&G monitors customers' comments on its brands and products on social media for new product ideas.
Leverage Big Data technology itself to differentiate product offerings	• Big Data technology itself can be leveraged to offer new or improve existing products. • Aviva Insurance in the UK leverages Big Data technology to track cars' GPS data to monitor drivers' driving patterns, and the information is used to assess appropriate discounts at subsequent insurance renewals.
Develop and roll out effective product launch strategy	• Pre-launch, firms can mine competitors', retailers' and distributors' website for market intelligence that are useful in developing their own product launch strategies. Post-launch, firms can mine online data to gauge the impact of their marketing activities across different channels and media, where the information can be used to assess marketing effectiveness and adjust marketing strategies accordingly to achieve set marketing objectives. • P&G monitors customers' comments after a product launch to improve its marketing effectiveness. Effective Measure, a firm providing online brand and advertising effectiveness measurement and targeting solutions, helps other firms conduct online measurement of their digital campaigns' effectiveness. STASIT, a social media marketing and technology firm, help other firms analyze social media data for campaign analysis, influencer analysis etc. Both Effective Measure and STASIT have a presence in Malaysia.

Sources: Aviva UK's website; McKinsey (2011b); Hanchard & Ramdas (2014); Anastasia, (2015)

4.2 Production

Firms can collect Big Data from their production processes, which can be analyzed using advanced analytics for valuable intelligence that can guide more effective production improvement actions. For example, Big Data can be used for predictive maintenance of equipment and yield improvement, as presented in Table 3. Table 4 provides two real world Big Data applications in production: case 1 is about predictive maintenance of equipment by a major provider of upstream oil field services; and case 2 is about yield improvement by Merck.[2]

Table 3: Big Data use cases in production

Use case	Description
Predictive maintenance of equipment	• A production workflow involves machines coordinated sophistically across pre-defined and precise steps, where one machine's malfunction may stop the whole production line, but premature maintenance will incur additional costs, so firms' challenge is to optimize their equipment maintenance schedules. • Firms can leverage sensors to collect massive amounts of data from their production equipment. The sensors can be programmed to collect data from temperature to speed to vibration level. Once a large volume of data has been accumulated, firms can mine the data to find wear patterns that cause equipment failures, and then proactively monitor real time data for preventive maintenance.
Yield improvement	• A 1% yield improvement can lower production costs significantly, thus a capability to analyze data for yield improvement can give manufacturers a tremendous competitive advantage. • McKinsey cited that a global chemical maker used more than 600 variables to model its plant's production yields under various operating conditions to identify improvement opportunities, and managed to increase the plant's earnings before interest and taxes by more than 50%.

Sources: Hortonworks' website; McKinsey (2014); WSJ. Custom Studios (2015); McKinsey (2015a)

Table 4: Two real world Big Data applications in production

Case 1: A major provider of upstream oil field services leveraged Big Data for predictive equipment maintenance
Problem • The firm managed huge base of costly equipment in fields in 80 countries. It had limited utilization data of its equipment. It would entail a lot of time and manual effort to collect and analyze data. The standard enterprise data warehouse that it had did not scale well, thus yielded incomplete results. Consequently it suffered high maintenance and replacement costs.

[2] McKinsey (2015a) outlines a 4-step approach to leverage Big Data for yield improvement: (1) use data visualization to identify initial patterns to prioritize data collection and analysis; (2) use correlation analyses to identify core determinants of process performance and form an initial hypothesis of root causes of yield drop and variability; (3) use significance testing to test initial hypothesis of root causes of yield drop and variability and focus on most statistically significant factors for further investigations; and (4) use artificial neural networks to model complex processes to quantify the impact of and optimal ranges for the identified parameters.

Application • It collected Big Data using sensors above and below ground. By mining accumulated data collected over a long period, it was able to identify wear patterns that could lead to equipment failures. It now monitors real time sensor data to proactively detect and respond to potential equipment failures. *Outcome* • Reduced manual time and effort in collecting and analyzing data, which facilitates equipment maintenance decisions, reduces expenses and improves safety.
Case 2: Merck improved production yield with Big Data
Problem • Vaccine production requires precise control of complex fermentation processes. Two batches of a vaccine produced using the same production process may nevertheless exhibit significant yield variances. • Merck used sensors to collect data on temperature trend, humidity level, flow rate, pressure, agitator speed etc. and it also kept records of equipment maintenance and calibration. Data for one vaccine was stored across 16 different systems. The data collected over 10 years amounted to about 1 billion records. It would be very time consuming to align and analyze the massive amounts of data from the disparate systems using the traditional spreadsheet-based approach. *Application* • It leveraged Big Data technology to integrate data from the 16 different systems and made more than 15 billion calculations and 5.5 million batch-to-batch comparisons. With access to all data at one single place it was able to perform virtual experiments to determine the best operating regions of the processes. This enhanced its ability to understand how the vaccine yield could be improved. *Outcome* • It found that certain characteristics in the fermentation phase of vaccine production were closely tied to yield in a final purification process and came up with a conclusive answer about yield variance within three months. This translated into a USD10 million profit impact, and according to McKinsey the yield increased by 50%.

Sources: Abyres Consultancy; Hortonwork; InformationWeek (2014)

4.3 Supply chain management

Typical functions of supply chain management are to work with suppliers to source parts and raw materials, manage inventory, get products out to the market etc. The main challenges include supply and demand mismatch, optimal supplier management, and optimal transportation and distribution management. Supply chain has a large cost component (Sanders & Meil, 2015). But it is also rich with data, which can be leveraged for insights to enhance efficiency and reduce cost. Table 5 provides some examples of Big Data use cases in supply chain management. Table 6 provides two real world Big Data applications in supply chain management: case 1 is about better demand forecasting by a global technology manufacturer; and case 2 is about distribution network optimization by a European consumer goods firm.

Table 5: Big Data use cases in supply chain management

Use case	Description
Improve demand forecast	• Firms can leverage supply chain Big Data to make better demand forecast to reduce supply chain burden. The Boston Consulting Group (BCG) estimates that firms that do a better job in predicting future demand can often cut inventory by 20% to 30%; increase average fill rate by 3% to 7%; and generate margin improvement by 1% to 2%.
Optimize supplier management	• Firms can develop models to generate performance predictions of various supplier mixes and to estimate supplier reliability. The models may take into account factors such as additional costs due to variations in speed with which different suppliers can deliver their goods; and switching costs due to reasons such as cancellation of long-term contract etc. This can provide firms better information when selecting their suppliers, such as to select suppliers that give the highest returns on investments in order to maximize profits.
Optimize transportation management	• Firms can leverage advanced analytics to dynamically analyze millions of real time data points from Radio Frequency Identification (RFID), Global Positioning Systems (GPS), road network, weather, accidents etc. to, for example, optimize delivery scheduling. BCG estimates that firms that use Big Data and advanced analytics to tackle such tough supply chain problem can identify saving opportunities equal to 15% to 20% of their transportation costs.
Optimize distribution network	• Firms' distribution networks tend to evolve into complex, inflexible and costly webs of warehouses, factories and distribution centers spreading across a huge territory as they expand over time. They can leverage their supply chain and sales data and advanced analytics for modelling and simulation to simplify their distribution networks. According to BCG, firms that leverage Big Data and advanced analytics to simply their distribution networks typically produce 10% to 20% savings in freight and warehousing costs, in addition to large savings in inventory.

Sources: BCG (2015); Martin (2015)

Table 6: Two real world Big Data applications in supply chain management

Case 1: A technology manufacturer leveraged Big Data for better demand forecasting
Problem • The firm's sales team gave overtly optimistic forecasts. This caused it ordered more inputs than was needed to ensure adequate supply. In addition, its suppliers in turn ordered too much from their own component suppliers. As a result, inventories increased throughout the whole value chain. *Application* • It analyzed more than 7 million data points including shipment records, historical

forecasting performance, and bill-of-material records to understand the causes of poor forecast performance. Root cause analysis helped identify the sources of the problem, which included the usual delays and operational breakdowns, but also more subtle but equally powerful factors such as misaligned incentives and an organizational structure with too many silos. *Outcome* • It expects to improve its forecast accuracy by 5% to 10%. The changes are expected to yield an increase in revenue, while lowering inventory, delivering better customer service, and reducing premium freight cost.
Case 2: A European consumer goods firm used Big Data to optimize its distribution network
Problem • It wanted to shift its country-based distribution network to a more efficient network spanning the European continent. It wanted to make use of its data in designing an optimal distribution network, but the volume and distribution of the data across different systems outstripped its existing IT capacity. *Application* • To identify an optimal distribution network, it used advanced analytics to model and analyze demand data, delivery data, inventory data etc. of its 30 brands spreading across more than 10 plants, and simulate multiple scenarios for delivery including full-truck load, direct-to-store delivery, and two-tier warehousing, as well as different transport rate structures based on load size and delivery direction. *Outcome* • It managed to reduce the number of warehouses from 80 to 20. Each of the remaining warehouses becomes larger and more efficient, as pooling demand across a smaller network of bigger warehouses decreases demand variability, thereby lowering required inventory level. It expects to reduce its operational expenditure by as much as 8%.

Source: BCG (2015)

4.4 Marketing

The scale of traditional marketing data (just think of point-of-sales data, responses to direct mail campaigns and coupon redemptions) is incomparable to Big Data today (just think of online purchase data, social media postings and geo-locational data). Marketers can leverage Big Data to address a variety of marketing challenges such as improving marketing mix, enhancing marketing efficiency, and understanding customer dynamics better. McKinsey found that firms that put data at the center of their marketing and sales decisions were able to improve their marketing's return on investment by 15% to 20% (McKinsey, 2013). Table 7 provides some examples of Big Data use cases in marketing.

Amazon analyzes its massive amounts of customer data to power its recommendation engine; and it once claimed that 35% of its sales was due to the recommendation engine (360i, 2012). As already mentioned earlier, Walmart has massive amounts of transactional, online and mobile data, and the firm leverages them to, among others, optimize the local assortments of its stores based on what their customers say on social media; recommend suitable products to Facebook users based on hobbies and interests of their networks; and provide targeted product recommendations to customers based on their purchase histories (Van Rijmenam, nd).

A Malaysian example of Big Data application in marketing is by Celcom. To stay competitive in the highly competitive and saturated market in which the firm operates, it has undergone major transformation to incorporate analytics into decision making in virtually every part of its business. For example, it leverages Big Data to deliver targeted promotions to customers based on their actual behaviors, such as by identifying prepaid customers with low weekend usage and encouraging them to reload on weekends by offering bonus airtime; and to optimize its product offerings, such as the Night Owl plan which offers customers a differentiated price in small hours when there is a lot of low-cost network capacity available. The transformation has led to, among others, 90% reduction in time spent on analytics, 80% reduction in new campaign launch time, 70% increase in campaign performance, and 98% employee engagement rating among its customer-facing employees. (IBM, 2014; MDEC, 2015)

Table 7: Big Data use cases in marketing

Use case	Description
Cross- and upselling	• Firms can analyze customer data to better understand their behaviors and preferences for more effective cross- and upselling. • Firms can also embed sensors into their products to proactively monitor usage by customers, and send alerts to their sales teams at the right time to contact customers for a product refresh, upgrade, or sell new products to them.
Identify new customer segment	• Marketers generally define their customer bases in terms of demographics, based on the belief that people who share similar characteristics are likely to behave in more or less the same way. Firms can mine Big Data to uncover new niche customers in terms of their actual behaviors and interests, rather than demographic characteristics.
Uncover new marketing opportunities	• Big Data can provide better insights into potential new product segments, product features, marketing opportunities etc. For example, Walmart found "cake pop" was trending on social media, so it swiftly introduced the product in its stores.
Drive more effective advertising	• Firms can leverage real time data on audiences' behaviors and product interests to deliver targeted one-to-one advertising messages. For example, an airline which knows that a customer is researching about a trip to Europe may deliver an advertising message to the customer that specifically refers to "Sales Fares to Europe".
Measure campaign impact	• Firms can mine Big Data such as clickstreams and social media conversations to learn how effective different channels

Use case	Description
	and media contribute towards campaign performance, and it can make use of the insights to improve its existing marketing mix or to make better budget allocation in future campaigns.
Reduce customer churn	• Firms can develop models and leverage Big Data to predict which customers are at risk of leaving, and obtain insights into what influences customer loyalty, and then execute targeted loyalty efforts to retain those at risk of leaving.
Monitor consumer sentiment	• Consumers are increasingly leveraging social media to voice out their grievances to make sure that they are heard. Firms can monitor social media to uncover consumer sentiment on their brands, products, campaigns etc. and then react accordingly and timely. • For example, before Greenpeace launched its anti-palm oil campaign against Nestlé in 2010, it had already been asking the firm for two years to stop procuring palm oil produced using unsustainable methods. When it started the campaign with a YouTube video, comments stormed social media. Without a proper digital strategy and being in panic mode, Nestlé reacted by asking YouTube to remove the video. This action created backlash and engendered more protests from the social media communities. If Nestlé had monitored social media, it would have understood how consumer sentiment had built up and be better prepared to react in this incident. After the incident, the firm created a Digital Acceleration Team (DAT) to monitor real time online conversations to help rebuild its reputation and for new insights to guide its future marketing campaigns.

Sources: The Economist (2010); Lonescu-Somers & Enders (2012); 360i (2012); Choudhury (2015)

5.0 Main challenges and recommendations for SMEs

Data is the new oil. SMEs who are not data-driven risk of being outcompeted by those who are. The complexity of today's business environment warrants SMEs not only to rely on gut feeling and past experience but adopt a more data-driven approach in their decision making. Big Data adoption is a strategy that SMEs can tap on to drive their competitiveness. However, there are challenges. The frequently cited Big Data challenges for SMEs are:

- SMEs lack Big Data;
- SMEs lack the big budget of big companies;
- There is a lack of Big Data talent; and
- It is a tough task to deliver big impact from Big Data, even for big, data-driven firms.

These challenges are in line with the findings in a research by UTAR on 132 Malaysian companies — mainly medium-sized and large companies — which show that majority (82%) of the firms are quite ready to take advantage and capture the huge benefits of Big Data, but they will need to make huge investment in technology, talent and culture (Wong, Chuah & Ong, 2015).

5.1 SMEs lack Big Data

To leverage Big Data, firms need to have Big Data. SMEs generally have small data sets, thus they may not need to invest in Big Data tools as traditional data tools may be sufficient to meet their needs.

Nevertheless, in today's Big Data environment, there are a lot of data out there. SMEs can leverage analytic products available in the market to gain access to external data. For example, Google Analytics provides access to web traffic data; IBM Watson Analytics provides access to Twitter; and MasterCard announced to provide merchants access to its massive amounts of payment data through IBM Watson Analytics on subscription basis starting from mid-year 2016 in several countries including Malaysia (Barbaschow, 2016).

Furthermore, with decreasing cost of data infrastructure, it has also become more affordable for SMEs to collect their own data. Data can be collected at multiple data points across business functional areas. For example, SME manufacturers can embed sensors into production equipment to collect data that are needed to improve production processes. SMEs are reported to have collected a wide range of data such transactional data, social media data, emails, geographic information system (GIS) data, web logs, sensor data, audio, and video. (IDC, 2012).

5.2 SMEs lack the big budget of big companies

Big Data may be an opportunity to drive business performance. However, insights can only be gained from data if firms have in place the means to leverage it (Springwise, 2015). To leverage Big Data, firms have to invest in Big Data capabilities, including investments in hardware and software, talent, as well as systems and processes to encourage a data-driven culture. Database technology and talent with the right skill sets do not come cheap. SMEs generally have a small IT team and a small data infrastructure in place, thus the initial investment required to develop Big Data capabilities can be considerable for them.

With limited capital and human resources, SMEs do not want customized technology products that take months to deploy, instead they want easy-to-understand, easy-to-use, and quick-to-deploy technology products which are also sensitive to their cash conservation requirements. Cloud-based technology products are marketed as the solutions for SMEs.

Infrastructure-as-a-Service (IaaS) offers low-cost data storage and data backup that gives businesses the computing resources they need to store their data, and Software-as-a-Service (SaaS) offers analytic engines that help businesses to analyze data stored on the cloud and provide analytic output to end users through a user-friendly interface. For these products, users only pay for the services they need, with prices charged on per-month, per-user and/or per-capacity-unit basis, and this reduces required upfront investment (see Table 8 and Table 9 for pricing examples of Oracle's cloud storage and IBM Watson Analytics respectively). They can also be deployed quickly as there is no need to set up any physical infrastructure, and are flexible as users can make incremental increase as their needs expand. In addition, SMEs can also use free open-source software such as Hadoop, Python and R.

Table 8: Oracle’s cloud storage pricing

Metered Services (Oracle IaaS Public Cloud Services)

Storage Capacity

per Month	Price	Metric
First TB/month	$0.0240	GB / Month
Next 49 TB/month	$0.0236	GB / Month
Next 450 TB/month	$0.0232	GB / Month
Next 500 TB/month	$0.0228	GB / Month
Next 4,000 TB/month	$0.0224	GB / Month
Over 5,000 TB/month	$0.0220	GB / Month

This bundled metered service allows you to set and activate the offerings currently available for Oracle Storage Cloud Service. Note that the purchase of the 'Pay as you go' service do not require any upfront payment; monthly invoice will be generated based on usage.

Source: Oracle’s website

Table 9: IBM Watson Analytics’ pricing

Free Edition	Plus	Professional
Use Watson Analytics for free and get access to cognitive, predictive and visual analytics. $0.00USD	All the analytics of the free version plus the ability to upload larger data sets. Starting at $34.50USD per month per user	Designed for enterprises, includes multi-user environment to collaborate and more data connectors. Starting at $92.00USD per month per user
1 user	1 user	1 or more users
100,000 rows per data set	1,000,000 rows per data set	10,000,000 rows per data set
50 columns per data set	256 columns per data set	500 columns per data set
500 MB of storage included	2 GB of storage included	100 GB of storage included
	Add more storage for an extra fee in increments of 10 GB	Add more storage for an extra fee in increments of 50 GB
Upload delimited files and Microsoft Excel files	Upload delimited files and Microsoft Excel files	Upload delimited files and Microsoft Excel files
	Access to more sources of data	Access to more sources of data
	Access to social data from Twitter	Access to social data from Twitter
		Connect to IBM Cognos report data
		Share data sets, refined data sets, explorations, predictions, and views

Source: IBM's website

SMEs can also hire external consultants instead of developing an in-house team. This will not only allow them to avoid the needs of making large upfront investment in data system but also to tap on external expertise. For example, they can hire consultants to help crunch online data for advertising effectiveness measurement, influencer analysis, campaign analysis, customer analysis, or to analyze store visitor traffic to assess staffing needs, staff performance, conversion ratios, floor plans, product display etc.

5.3 There is a lack of Big Data talent

To benefit from Big Data, firms need talent with the necessary skill sets. There is a shortage of Big Data talent in the market. It is estimated that the US alone may face a shortage of 140 000 to 190 000 people with deep analytical skills as well as 1.5 million managers and analysts with know-how of Big Data analysis by 2018 (McKinsey, 2011a). As a late comer in Big Data, Malaysia too lacks Big Data talent. According to Malaysia's Minister of Higher Education, there were 4 000 Big Data scientists in the country in 2014 and an additional 12 000 will be needed in the next five years (The Star Online, 2015). SMEs are generally seen as less attractive employers compared with big companies due to perception of limited career opportunities and less competitive remuneration. It may therefore be less attractive for SMEs to adopt Big Data since it will be useless to invest in the necessary infrastructure if no one is there to make sense of the data and to make the insights actionable.

In Malaysia, MDeC is working with the industry and the education and training sector to develop Big Data talent. Nevertheless, until a local pool of skilled talent is grown, it will continue be necessary for firms that lack their own Big Data talent to outsource work to talent elsewhere. Some firms have turned to Kaggle for help, an online platform that matches data-related requests with data scientists from around the world at a far more affordable cost, which may significantly reduce firms' financial commitments of developing the necessary capabilities in-house.

5.4 It is a tough task to deliver big impact from Big Data, even for big, data-driven firms

To benefit from Big Data, firms need to transform insights into value, and this has been difficult even for big, data-driven firms. According to McKinsey, few firms have achieved what it calls "big impact from Big Data" (McKinsey, 2015b). When the consulting firm asked a group of analytics leaders from major firms what degree of revenue or cost improvement they had achieved through the use of Big Data and advanced analytics, 3/4 of them said it was less than 1%. McKinsey pointed out that although interesting customer insights may be obtained from social media analysis, they may not be actionable or able to influence business results in a meaningful way, and furthermore, frontline managers and business users often see Big Data tools like black boxes so they do not act on the Big Data insights.

To increase chances of Big Data success, firms embarking on Big Data projects are advised to develop a data-driven culture in their organizations; to carefully evaluate the business case of their Big Data projects; and to clearly define the specific outcomes, i.e. to think of what to ask of the data, how the firm will react to the answers, and what are the actionable operational measures (IDA, 2012).

6.0 Conclusion

Big Data can provide richer, more granular and faster insights, so it represents an opportunity for firms to create a unique competitive advantage to win in the ever more competitive marketplace. Many firms have leveraged Big Data to solve a wide range of business problems and achieved considerable success. A number of research have also found that firms who are data-driven perform better than those who are not. But leveraging Big Data entails considerable investments in technology, talent and culture. Nevertheless, in today's Big Data environment, there are increasing availability of cheaper and flexible cloud-based technology solutions, which are particularly suitable for SMEs. Talent is a major obstacle because Big Data talent is scarce in the market. Firms can engage established external consultants or turn to the Internet to find freelance data scientists at a far more affordable rate. Big Data can provide interesting insights but they have to be acted on and transformed into value to generate returns, and there is no guarantee of success. To increase chances of success, firms should develop a data-driven culture throughout their organizations.

References

360i. (2012). *The CMO's Guide to Big Data.* Retrieved from https://360i.com/pov/big-data/

Anantharam, D. (February, 2013). Big Data, Big Deal. Retrieved from Big data, big deal: https://blogs.msdn.microsoft.com/data__knowledge__intelligence/2013/02/18/big-data-big-deal/

Anastasia. (February, 2015). Big Data and New Product Development. Retrieved from https://www.cleverism.com/big-data-new-product-development/

BCG. (29 May, 2013). *How to Get Started with Big Data.* Retrieved from https://www.bcgperspectives.com/content/articles/it_strategy_retail_how_to_get_started_with_big_data/

BCG. (January, 2015). Making Big Data Work: Supply Chain Management. Retrieved from https://www.bcgperspectives.com/content/articles/technology_making_big_data_work_supply_chain_management/

Choudhury, S. (September, 2015). How Nestlé Uses Data Visualization for Social Media Monitoring and Engagement? Retrieved from http://www.fusioncharts.com/blog/2014/08/how-nestle-uses-data-visualization-for-social-media-monitoring-and-engagement/

Dasgupta, A. (2015). Big Data and Crowdsourcing. Retrieved from http://gtupgcenters.edu.in/News/Report%20of%20Expert%20Lecture%20Big%20Data%20and%20Crowdsourcing%20by%20Professor%20Anirban%20Dasgupta.pdf

DHL. (2013). *Big Data in Logistics: A DHL Perspective on How to Move Beyond the Hype.* DHL Customer Solutions & Innovation. Retrieved from http://www.delivering-tomorrow.de/wp-content/uploads/2015/08/CSI_Studie_BIG_DATA_FINAL-ONLINE.pdf

Digital News Asia. (January, 2014). Next Step for Digital Malaysia: The 'DM354' Roadmap. Retrieved from https://www.digitalnewsasia.com/digital-economy/next-step-for-digital-malaysia-the-dm354-roadmap

EITO. (2013). *Big Data in Europe: Evolution and Revolution.* Retrieved from http://leanbi.ch/wp-content/uploads/studien/LeanBI_EITO_Big_Data_in_Europe.pdf

Hanchard , S., & Ramdas, T. (2014). *Big Data in Malaysia: Emerging Sector Profile 2014.* Retrieved from http://bigdataanalytics.my/downloads/big-data-malaysia-emerging-sector-profile-2014.pdf

Hanmark, K. (May, 2015). Redefine Big Data for Your Business. Retrieved from http://data-informed.com/redefine-big-data-for-your-business/

Hoskins, M. (February, 2015). How Graph Analytics Deliver Deeper Understanding. Retrieved from http://www.infoworld.com/article/2877489/big-data/how-graph-analytics-delivers-deeper-understanding-of-complex-data.html

IBM. (October, 2013). Data Deluge Needs Smarter Storage. Retrieved from Data deluge needs smarter storage: http://forbesindia.com/blog/technology/data-deluge-needs-smarter-storage/

IBM. (2014). *Celcom Axiata: Using Analytics to Drive Customer Engagement and Market Leadership.* Retrieved from https://www.ibm.com/smarterplanet/global/files/us__en_us__leadership__odc03230_00_hr.pdf

IDA. (2012). *The IDA Infocomm Technology Roadmap (ITR) 2012.* Retrieved from https://www.ida.gov.sg/Tech-Scene-News/Technology/Technology-Roadmap

IDC. (October , 2012). *How Small and Midsize Enterprises Can Sharpen Performance with Next Generation Business Intelligence and Analytics.* Retrieved from http://mwaintel.com/wp-content/uploads/2013/11/How_Small_and_Midsize_Enterprises_Can_Sharpen_Performance_with_Next-Generation_Business_Intelligence_and_Analytics.pdf

IDC. (April, 2014). *The Digital Universe of Opportunities: Rich Data and the Increasing Value of the Internet of Things.* Retrieved from http://www.emc.com/leadership/digital-universe/2014iview/executive-summary.htm

IDC. (November, 2015). New IDC Forecast Sees Worldwide Big Data Technology and Services Market Growing to $48.6 Billion in 2019, Driven by Wide Adoption Across Industries. Retrieved from http://www.idc.com/getdoc.jsp?containerId=prUS40560115

InformationWeek. (February, 2014). Merck Optimizes Manufacturing With Big Data Analytics. Retrieved from http://www.informationweek.com/strategic-cio/executive-insights-and-innovation/merck-optimizes-manufacturing-with-big-data-analytics/d/d-id/1127901

Inside BIGDATA. (June, 2015). Text Analytics: The Next Generation of Big Data. Retrieved from Text Analytics: The next generation of Big Data: http://insidebigdata.com/2015/06/05/text-analytics-the-next-generation-of-big-data/

Kepes, B. (September, 2015). Using Graph Database Technology to Unravel Banking Fraud. Retrieved from http://www.computerworld.com/article/2979654/data-analytics/using-graph-database-technology-to-unravel-banking-fraud.html

Lonescu-Somers, A., & Enders, A. (December, 2012). How Nestlé Dealt with a Social Media Campaign Against It: Engaging with Critics Was an Effective Response. *Financial Times*. Retrieved from https://next.ft.com/content/90dbff8a-3aea-11e2-b3f0-00144feabdc0

Loshin, D. (October, 2005). Simple Semi-Structured Data. Retrieved from Simple semi-structured data: http://www.b-eye-network.com/view/1761

Malay Mail. (January, 2015). Big Data: Malaysia Takes ‘Small but Significant’ First Step. Retrieved from http://www.themalaymailonline.com/tech-gadgets/article/big-data-malaysia-takes-small-but-significant-first-step

Martin. (January, 2015). How to Optimize Supply Chain Management with Big Data. Retrieved from https://www.cleverism.com/how-to-optimize-supply-chain-management-big-data/

McAfee, A., & Brynjolfsson, E. (October, 2012). Big Data: The Management Revolution. *Harvard Business Review*. Retrieved from https://hbr.org/2012/10/big-data-the-management-revolution/ar

McKinsey. (May, 2011a). *Big Data: The Next Frontier for Innovation, Competition, and Productivity.* Retrieved from http://www.mckinsey.com/business-functions/business-technology/our-insights/big-data-the-next-frontier-for-innovation

McKinsey. (November, 2011b). Inside P&G’s Digital Revolution. *McKinsey Quarterly*. Retrieved from http://www.mckinsey.com/industries/consumer-packaged-goods/our-insights/inside-p-and-ampgs-digital-revolution

McKinsey. (2013). Smart Analytics: How Marketing Drives Short-Term and Long-Term Growth. In *Big Data, Analytics And The Future Of Marketing And Sales.*

McKinsey. (2014). How Big Data Can Improve Manufacturing. Retrieved from http://www.mckinsey.com/business-functions/operations/our-insights/how-big-data-can-improve-manufacturing

McKinsey. (February, 2015a). Taming Manufacturing Complexity with Advanced Analytics. *McKinsey Quarterly*. Retrieved from http://www.mckinsey.com/business-functions/operations/our-insights/taming-manufacturing-complexity-with-advanced-analytics

McKinsey. (January, 2015b). Getting Big Impact from Big Data. *McKinsey Quarterly*. Retrieved from http://www.mckinsey.com/business-functions/business-technology/our-insights/getting-big-impact-from-big-data

MDEC. (2015). *Exposing the Promise of Big Data Analytics.* Retrieved from http://bigdataanalytics.my/wp-content/uploads/2015/11/BDA-case-studies-in-Malaysia-lowres.pdf

Nadarajah, P. (2015). Valuing Big Data assets. *Accountants Today*(November/December 2015).

Neo Technology. (2014). *Powering Recommendations with a Graph Database.* Retrieved from http://info.neotechnology.com/rs/neotechnology/images/GraphDB_Recommendations_EN.pdf

Neo Technology. (nd). Walmart Optimizes Customer Experience with Real-time Recommendations. Retrieved from http://neo4j.com/case-studies/walmart/

Pafka, S. (November, 2015). Big RAM is Eating Big Data – Size of Datasets Used for Analytics. Retrieved from http://www.kdnuggets.com/2015/11/big-ram-big-data-size-datasets.html

Quantum. (2015). *Introducing DXi with Rocket Arkivio Autostor: The Smart Way to Manage Your (Unstructured) Data.* Retrieved from https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&cad=rja&uact=8&ved=0ahUKEwjCpb2GrtPLAhVLG44KHXniCOcQFggcMAA&url=http%3A%2F%2Fwww.quantum.com%2Fiqdoc%2Fdoc.aspx%3Fid%3D11004&usg=AFQjCNGm44-JnzX2aWT0XPV8v6SA4kEIyA&sig2=IBHwtXyWkXdc1KbwgyUEIA

Ross, J. W., Beath, C. M., & Quaadgras, A. (December, 2013). You May Not Need Big Data After All. *Harvard Business Review*. Retrieved from https://hbr.org/2013/12/you-may-not-need-big-data-after-all

Sanders, R., & Meil, J. (November, 2015). Big Data: The Latest Rage in Supply Chain Management. Retrieved from http://ww2.cfo.com/supply-chain/2015/11/big-data-latest-rage-supply-chain-management/

SAP AG. (2012). Harnessing the Power of Big Data in Real Time through In-Memory Technology and Analytics. *The Global Information Technology Report*. Retrieved from http://www3.weforum.org/docs/Global_IT_Report_2012.pdf

SAS. (nd). How Walmart Makes Data Work for Its Customers. Retrieved from http://www.sas.com/en_us/insights/articles/analytics/how-walmart-makes-data-work-for-its-customers.html

Siegler, M. (August, 2010). Eric Schmidt: Every 2 Days We Create As Much Information As We Did Up To 2003. Retrieved from https://techcrunch.com/2010/08/04/schmidt-data/

SimaFore. (2011). *Affordable Analytics for Small and Medium Business: Cost Modelling with Predictive Analytics Improves Forecast Capability of a Mid-Sized Manufacturing Company.* Retrieved from http://cdn2.hubspot.net/hub/64283/pdfs/simafore-whitepaper-cost-modeling.pdf

Spire. (January, 2016). Deciphering Unstructured Data with Contextual Intelligence. Retrieved from http://www.spiretechnologies.com/deciphering-unstructured-data-with-contextual-intelligence/

Springwise. (August, 2015). The Age of 'Big Data', and What It Means Today. Retrieved from http://www.springwise.com/age-big-data-means-today/

The Economist. (June, 2010). The Campaign Against Palm Oil: The Other Oil Spill. Retrieved from http://www.economist.com/node/16423833

The Star Online. (December, 2014). MDeC: Big Data Analytics A Key Focus for ICT Development in Malaysia. Retrieved from http://www.thestar.com.my/tech/tech-news/2014/12/16/mdec-big-data-analytics-a-key-focus-for-ict-development-in-malaysia/

The Star Online. (September, 2015). Idris: We Need 12,000 Big Data Scientists for Global Economy. Retrieved from http://www.thestar.com.my/news/nation/2015/09/04/idris-we-need-12000-big-data-scientists-for-global-economy/

The Sun Daily. (April, 2015). MDeC Aims to Make Malaysia Big Data Analytics Hub in Asean. Retrieved from http://www.thesundaily.my/news/1391560

Van Rijmenam, M. (May, 2014). Why UPS Spends Over $ 1 Billion on Big Data Annually. Retrieved from https://datafloq.com/read/ups-spends-1-billion-big-data-annually/273

Van Rijmenam, M. (nd). Walmart Is Making Big Data Part Of Its DNA. Retrieved from https://datafloq.com/read/walmart-making-big-data-part-dna/509

Wagner, J. (July, 2014). Graph Technology Provides Valuable Insights From Interconnected Data. Retrieved from http://www.programmableweb.com/news/graph-technology-provides-valuable-insights-interconnected-data/analysis/2014/07/21

Wong, K., Chuah, M., & Ong, S. (2015). Are Malaysian Companies Ready for the Big Data Economy? A Business Intelligence Model Approach. *International Conference on Accounting Studies (ICAS) 2015.* Retrieved from http://www.icas.my/download/icas_2015/113-ICAS2015_W.%20K.%20Luen,%20C.%20M.%20Hooi,%20O.%20S.%20Fook.pdf

WSJ. Custom Studios. (2015). *Manufacturing and the Search for more Intelligence.*

SOCIO-ECONOMIC RESEARCH CENTRE (SERC)
SERC SDN BHD (918837-W)
6th Floor, Wisma Chinese Chamber,
No. 258, Jalan Ampang, 50450 Kuala Lumpur, Malaysia.
Tel : (603) 4260 3116 / 3119
Fax : (603) 4260 3118
Email : serc@acccimserc.com
Website : http://www.acccimserc.com