



International Journal of Operations & Production Management

Making sense of Big Data - can it transform operations management?

Olga Matthias Ian Fouweather Ian Gregory Andy Vernon

Article information:

To cite this document:

Olga Matthias Ian Fouweather Ian Gregory Andy Vernon , (2017), "Making sense of Big Data – can it transform operations management? ", International Journal of Operations & Production Management, Vol. 37 Iss 1 pp. 37 - 55

Permanent link to this document:

<http://dx.doi.org/10.1108/IJOPM-02-2015-0084>

Downloaded on: 07 January 2017, At: 11:56 (PT)

References: this document contains references to 75 other documents.

To copy this document: permissions@emeraldinsight.com

The fulltext of this document has been downloaded 16 times since 2017*

Users who downloaded this article also downloaded:

(2017), "Role of social media in retail network operations and marketing to enhance customer satisfaction", International Journal of Operations & Production Management, Vol. 37 Iss 1 pp. 105-123 <http://dx.doi.org/10.1108/IJOPM-03-2015-0153>

(2017), "Challenges and opportunities of digital information at the intersection of Big Data Analytics and supply chain management", International Journal of Operations & Production Management, Vol. 37 Iss 1 pp. 10-36 <http://dx.doi.org/10.1108/IJOPM-02-2015-0078>



Access to this document was granted through an Emerald subscription provided by emerald-srm:549055 []

For Authors

If you would like to write for this, or any other Emerald publication, then please use our Emerald for Authors service information about how to choose which publication to write for and submission guidelines are available for all. Please visit www.emeraldinsight.com/authors for more information.

About Emerald www.emeraldinsight.com

Emerald is a global publisher linking research and practice to the benefit of society. The company manages a portfolio of more than 290 journals and over 2,350 books and book series volumes, as well as providing an extensive range of online products and additional customer resources and services.

Emerald is both COUNTER 4 and TRANSFER compliant. The organization is a partner of the Committee on Publication Ethics (COPE) and also works with Portico and the LOCKSS initiative for digital archive preservation.

*Related content and download information correct at time of download.

Making sense of Big Data – can it transform operations management?

Making sense
of BD

Olga Matthias, Ian Fouweather and Ian Gregory
*Department of Operations and Information Management,
University of Bradford School of Management, Bradford, UK, and*
Andy Vernon
PA Consulting Group, Manchester, UK

37

Received 15 February 2015
Revised 17 September 2015
13 December 2015
Accepted 16 December 2015

Abstract

Purpose – The purpose of this paper is to focus on the application and exploitation of Big Data (BD) to create competitive advantage. It presents a framework of application areas, and how they help the understanding of targeting and scoping specific areas for sustainable improvement. Empirical evidence demonstrates the application of BD in practice and tests the framework.

Design/methodology/approach – An exploratory approach is adopted to the secondary research which examines vendors' offerings. The empirical research used the case study method.

Findings – The findings indicate that there is opportunity to create sustainable competitive advantage through the application of BD. However there are social, technological and human consequences that are only now beginning to emerge which need to be addressed if true long-term advantage is to be achieved.

Research limitations/implications – The research develops a framework and tests it only in two dimensions. This should be expanded. The vendor analysis limitations lie within the nature of the information available and the difficulties in mitigating against bias.

Practical implications – The suggested framework can help academics and managers to identify areas of opportunity to do so, setting new levels of performance and new agendas for business.

Originality/value – This work contributes to service operations management, building on Kranzberg (1986) and the impact of technology and on Fosso Wamba *et al.* (2015) by developing a systems application framework to further understanding of BD from a practical perspective to extend their research taxonomy insights. The case studies demonstrate how the use of BD enhances operational performance.

Keywords Big Data, Operational performance, Methodology for operations management, Business performance, Data analysis

Paper type Research paper

1. Introduction

Technology is a major source of change in today's business world. As internet and mobile technologies grow, all aspects of our lives are being transformed. Enterprises generated and stored an estimated seven exabytes (billions of gigabytes) of data in 2010, whilst consumers generated an additional six exabytes (Manyika *et al.*, 2011). The IDC (2012) estimates that by 2020 this will increase to 40 exabytes. Exponential technological growth can bring social and organisational challenges. As Kranzberg (1986, p. 545) notes:

Technology is neither good nor bad; nor is it neutral [...] technical developments frequently have environmental, social and human consequences that go far beyond the immediate purposes of the technical devices and practices themselves.

Technology companies (typified by Google, Microsoft, IBM, Oracle and SAP) and commentators such as Bell (2013) and Porter and Heppelmann (2014) promote the need for businesses to exploit data streams to create competitive advantage. The OECD (2013, p. 4) suggested that:

The exploitation of data promises to create added value in a variety of operations ranging from optimising the value chain and manufacturing production to more efficient use of labour and better customer relationships.



If organisations can harness internally or externally generated data, their operational capabilities could be transformed. Whether by providing greater volumes of reliable and timely information for decision makers, or by automating decision-making processes, there is a collective assumption that Big Data (BD) will benefit organisations, individuals and society (Bughin *et al.*, 2011; Barton and Court, 2012). Computer applications capable of analysing huge data sets are becoming readily available (Fisher *et al.*, 2012). Opresnik and Taisch (2015) highlight that the challenge for organisations is to develop strategies that exploit BD to generate added value. Others question is whether BD will ultimately deliver what it promises (Croxall, 2014; Lury, 2013). Gandomi and Haider (2015) observe that there has been little critical discourse, or empirical academic research, into BD and how it might be harnessed.

The research described in this paper arose from a curiosity about the possible benefits of BD and the lack of clarity about exactly how amassed data can be helpful to organisations. The paper empirically examines current application of BD to create competitive advantage, using current literature and secondary research, to evaluate applications. This broad assessment of systems' solutions purchasing enables the development of a framework indicating applications usage. The framework is then evaluated using as case studies two UK retail organisations which sought to exploit available BD.

The paper adds to a small but growing body of operations management literature, alongside work such as Demirkan and Delen (2013), Fosso Wamba *et al.* (2015) and Huang and Handfield (2015). It makes a contribution to knowledge by developing a framework detailing how BD applications are currently being exploited. It contributes to practice by helping organisations develop their operational strategy to best exploit the available data to engender sustainable improvement.

2. Literature review

There is considerable hype around the term BD (Gandomi and Haider, 2015; George *et al.*, 2014; Gartner, 2013, 2015). It has become endemic since the emergence of the term in the mid-1990s (Diebold, 2012). Technology companies, whether providers of ERP systems, CRM software or business analytics have been promoting BD for over a decade. In 2012 McAfee and Brynjolfsson (2012), writing in the *Harvard Business Review*, said BD represented a "revolution in management". Brown *et al.* (2011) asserted that it would "transform business processes and alter corporate ecosystems". Manyika *et al.* (2011) described it as "the next frontier for innovation, competition and productivity".

The rhetoric of these messages seems to be shaping the expectations of academic communities. Fosso Wamba *et al.* (2015) provided a comprehensive review of existent literature on BD. They concluded that the majority of publications focussed on BD technologies and access to data. To broaden the understanding of the role BD have in capturing business value they developed a general taxonomy from this review. Empirical research by Wieland *et al.* (2014) suggests that within the field of supply chain management academics expect BD to be "the hot topic" for the next five years. They indicate that leading researchers acknowledged a degree of scepticism that it could be the latest in a long line of management fashions and fads. This view is supported by Madsen and Stenheim (2013).

2.1 The antecedents of BD

The ability to generate vast streams of data has increased as the ability to rapidly process that data has increased. Moore's (1965) Law stated that the processing power of computers doubled every 18 months. Kryder's Law says that digital storage is increasing at a similar rate to data processing power (Esener *et al.*, 1999). Hruska (2013) questions whether Moore's Law remains valid and suggests that current storage technologies are reaching their limit.

In the 1990s systems engineers acknowledged that greater volumes of data, generated at higher rates, do not automatically lead to more information and knowledge. Without the ability to analyse and understand the data being generated, possessing more data may actually be a hindrance. Unless data can be transformed into information that facilitates good decision making and enhances operational performance, it has no purpose (Jifa and Lingling, 2014). Unlocking the knowledge within the data remains the central challenge.

The link between data and information has been explored by systems theorists for decades and the insights provided have shaped operations management for many years (Carvajal, 1992). It is through the “application of data and information” that knowledge is created (Ackoff, 1989). Knowledge is the ability to use information within a particular context. Shedroff’s (1994) DIKW hierarchy links data, information, knowledge and wisdom. It provides a theoretical pyramid built on a foundation of data with each successive layer resting on the one beneath. Focus has generally been on the three lower levels of the hierarchy (Rowley, 2007).

Information is:

[...] organised or structured data, which has been processed in such a way that the information has relevance for a specific purpose or context, and is therefore meaningful, valuable, useful and relevant (Rowley, 2007).

Put more practically, by processing data appropriately it becomes “useful for decisions and or action” (Liew, 2007). If information is “know what”, knowledge is “know why”. Knowledge permits decision making, enabling the selection of a particular action from a range of possibilities and leads to the “know how” to improve operational performance and create competitive advantage. With bigger (and by implication better) data comes the promise of enhanced knowledge and decision making.

2.2 The emergence of BD

There is no universally accepted definition of BD. McKinsey Global Institute’s study (2011) suggests that: “[...] data sets whose size is beyond the ability of typical database software tools to capture, store, manage and analyse”. The Gartner (2014) definition is “Data assets that require innovative forms of information processing for enhanced insight and decision-making”. Both seem to imply that the challenge of BD is scale, or volume, omitting any mention of the human aspect or the need to understand how meaning is created from BD, despite Kranzberg (1986) identifying this issue some 30 years earlier.

Although the origins of the term BD are obscure, there is agreement about when it appeared in academic literature (O’Leary, 2013; Jifa and Lingling, 2014). Cox and Ellsworth (1997) presented a paper at the IEEE’s 8th Conference on Visualization. Their narrative presented BD as a problem for systems engineers pondering how large data sets that exceeded available memory capacity could be managed. Subsequently Laney (2001) published a research paper in which he predicted that with the rise of e-commerce, enterprises would need to manage ever larger data sets and the challenge would be in terms of three dimensions: volume, velocity and variety. The 3Vs have become central to understanding BD. Essentially BD refers to high velocity, variable and complex data requiring advanced technologies and techniques to capture, store, distribute and manage for subsequently useful analysis. Over time the 3Vs have been extended with the addition of additional dimensions: variability, veracity and value. However, the 3Vs remain core for any understanding. There has also been a shift away from considering the challenges of BD to a more speculative approach that looks towards the future benefits of BD.

Traditional business analytics have focussed on internally created data. An essential element of the discourse on BD is social data (Chae, 2015). The proliferation of smart technologies has created a constant stream of data about individuals that can be utilised.

Additionally, social media offers a real-time window into people's opinions, wants and needs, not previously accessible. Within operations management the focus has been on internal data related directly to operations and processes whilst considering the behaviour and expectations of customers. BD offers new ways to understand the external environment (Dubey *et al.*, 2016).

A complication in utilising social media-generated data which has had little attention is how customers are affected. With data and information at their fingertips, customer expectations regarding the service they receive change. Customers create their own choices and their own relationship with a business by the online searches they make. They expect interaction through a channel of their choice at a time convenient to them with the same experience regardless of channel or device used. Inevitably changes in managing customer-facing processes are required. Most importantly perhaps, the dynamics of trust and co-production are altered. Little research exists on this in mainstream operations management.

2.3 BD research

Gandomi and Haider (2015) and Fosso Wamba *et al.*'s (2015) work shows the rapid growth in the use of the term BD in publications and corroborates Wieland *et al.*'s (2014) description of it as a "hot topic". As more articles emerge, the significance of the term increases and a positive feedback loop creates greater usage, further amplifying the significance collectively attached to the term, as illustrated by the twitter storms phenomenon (Seegerberg and Bennett, 2011).

Yet research on data analytics and operations improvement is not new. For instance, Bell (1999, p. 307) cites a FedEx executive referring to the RFID transmitter, indicating his organisation had "succeeded by applying scientific methods to its operations" utilising advanced analytics. The RFID transmitter is a smart object still referred to in mainstream operations management papers, and now in BD (Ilic *et al.*, 2010; Zelbst *et al.*, 2012; Meyer *et al.*, 2014; Lee and Özer, 2007; Zhong *et al.*, 2016). In their recent work Chongwatpol and Chan (2015) describe a case study which used a large dynamic data set to enhance operational decision making and increase effectiveness. Their research shows how data analytics can be used to find alternative ways of assessing business issues. They did not use the term BD.

Thus, whilst empirical BD research is not yet widespread in the operations management community, the technologies and analytic techniques required to deliver BDs promised value are well-researched.

2.4 The big idea of BD

In 2013 almost one billion smartphones were sold (Gartner, 2014) each capable of creating and collecting masses of data. In an operational environment smart machines are commonplace; new equipment comes with an array of sensors and data trackers that produce and store endless data (Lucke *et al.*, 2008; Meyer *et al.*, 2014; Zelbst *et al.*, 2012). Technology companies are keen to develop more applications that will generate even more data. Customer transactions and electronic feedback can be processed to capture demand and levels of satisfaction. Internally myriad smart technologies generate instantaneous feedback on the status and performance of internal resources. The claim is that processes can be controlled more effectively and better decisions to exploit opportunities and solve problems taken because of the information generated by new technology.

The Google Flu Trends (GFT) project reveals just how such claims can prove disappointing (Lazer *et al.*, 2014). The project's goal was to develop a means of identifying the emergence of flu so that health resources could be mobilised to treat the illness and prevent an epidemic. Using readily available data from Google's search engine, data on the frequency of "flu" searches were collected. A sophisticated analytical tool was developed to

extrapolate from the search data to predict future rates of flu. After apparent early success the model was found to consistently over-inflate future occurrences and was less accurate than existing “small data” strategies that utilised data on confirmed cases of flu. Their work highlights how the absence of criticality and rigour leads to flawed results: “quantity of data does not mean that one can ignore foundational issues of measurement and construct validity and reliability and dependencies among data” (Lazer *et al.*, 2014).

BD usefulness is constrained by the ability of the researcher to ask the right question, the same constraint as in 1943 when Abraham Wald was analysing appropriate positioning of armour on fighting aircraft. The predominant thinking was to armour those areas where it was obvious that aircraft had sustained damage. Wald’s wondered where the damage might be in those aircraft that had not returned. The only data he had were the analysis of bullet holes on planes that had returned. He successfully developed an assessment methodology used in subsequent wars (notably Korea and Vietnam) to minimise enemy-inflicted damage (1980).

In the case of the conclusions drawn from the GFT data, the absence of criticality, curiosity and validity resulted in erroneous projections. There is no way of knowing who searches for data on flu, or, equally importantly, why. The data collection only demonstrates the population that have an interest in flu – people with symptoms, and/or their friends, family, work colleagues. Data mining (DM) and analytic techniques may reveal interesting patterns hidden within large data sets. Without an understanding of what the patterns reveal, the information may have little value or, worse, lead to incorrect conclusions. As GFT shows, searching for patterns is complex. Lazer *et al.*’s message is clear: it is the insight and understanding of those involved in the analysis of data (whether big or small), that is central in creating value. The example of Wald remains an important reminder.

Uncritical analysis of poorly understood data sets does not generate knowledge. As Li *et al.* (2016) acknowledge, despite the increasing availability of data, how BD can be used to support decision making is “an enormous challenge”. Whatever the size of the data set, it needs appropriate analysis to create useful information that reveals what is significant within the data. “Not all information is useful for improving our understanding and judgements” (Saaty, 2008) and too much information can create uncertainty, hindering decision making. A certain amount of inference is required, and possibly the use of Bayesian analysis in support. Only then the available data will facilitate effective decision making that can add value to operations and enhance performance.

Piccoli and Pigni’s (2013) work illustrates how digital data streams (DDS) can be used within an operational setting to replace routine decision-making activities with predetermined (mindless) responses. This proven technology creates far higher levels of automation than previously possible, but the authors stress that DDS is not BD. Rather, it is an example of closed-loop feedback control. Research by Sætre *et al.* (2003) found that over-reliance on technology can create mindless responses if individuals fail to contextualise information, more likely to reduce operational performance than improve it. If BD are to be exploited to assist in decision making it needs to be analysed and utilised with care.

Bisel *et al.* (2014) indicate that there is an assumption that the size of a data set is seen by many researchers and analysts as a proxy for quality, and yet as Lazer *et al.* (2014) stress, quantity is not a substitute for quality. The volume and variety of BD make assessing its veracity challenging. Whilst statistical tools can be used to filter out erroneous and missing data, establishing the veracity of a large data set is not straightforward. Chongwatpol and Chan (2015) illustrate just how much work is required to ensure veracity of data and the information generated from large dynamic data sets.

Significantly, this is not a matter of technology, but one of skills. What new skills are required is something that writers do not necessarily agree upon (Miller, 2014). The OECD (2013, p. 29) acknowledges that an appropriate mix of advanced ICT, statistics and specific sector skills are required. Fawcett and Waller (2014) similarly recognise the need for both

data and domain skills, but stress that it is the ability to apply technical skills that is important. Manyika *et al.* (2011) highlighted that organisations do not have the talent to derive insights from BD. When presented with a complex and ever-changing stream of data and information, the ability to think creatively, grasp the situation and act accordingly may require a diverse range of attributes, such as intelligence, intuition, imagination and creativity, not always recognised or valued in the workplace. Despite advances in AI, technology cannot provide these attributes. Mainstream media (Philipson, 2014) recently reported that the British Intelligence Service (GCHQ) recognised this and employed more than 100 dyslexic and dyspraxic analysts because of their skills in identifying patterns and ability to analyse complex data.

Analysing BD is complex for a variety of reasons. Traditionally data analysis has considered discrete data that can be handled using well-established and sophisticated quantitative techniques, such as data feeds from sensors. Processing such data is straightforward, and can easily be automated. The data generated through social media present a greater challenge. Data are unstructured and come in a range of formats, often with multimedia content and threads of previous textual dialogues embedded within it. As Bisel *et al.* (2014) note, analysing such qualitative textual data requires specific skills. Analysts are working on tools to achieve this, but the ability of software to analyse text remains rather limited (Chen *et al.*, 2012). One exception is the work of Bollen *et al.* (2011). They successfully analysed the “mood” of Twitter feeds to predict movement in the Dow Jones Industrial Average, identifying an accuracy of 86.7 per cent. Yet Tinati *et al.* (2014) observe that the current forms of analysis of rich unstructured qualitative data is limited to classifying, linking and revealing distributions of words, reducing it to little more than a word-count survey. Whilst these are essential processes of data reduction which allow analysts to sift through large data sets quickly, they leave much unexplored. The role of the analyst is central to the process of interpreting the data, as demonstrated by Wald. Housley *et al.* (2014) detail work that is being undertaken to enhance capabilities in this area, but sophisticated tools and techniques have yet to be established.

There is an additional dimension to complexity that is often overlooked when considering the challenges of BD. Jifa and Lingling (2014) highlight the problems of working with open, complex, giant (OCG) systems. Their insight reveals that the internet creates an interconnected system, and that what is being harnessed is not simply more data about an existing system, it is data which is further complicating the system. It is not a closed system of the sort created within operations; it is a system constantly changing, with positive and negative feedback loops interacting dynamically, which can thereby simultaneously create stability and chaos. This confirms GCHQ’s conclusion that using data from open complex systems requires a different set of skills. These are the sort of skills dyslexic pattern-spotters have been forced to develop to survive in a text-orientated world. Their ability to spot oddities in patterns, which must have been useful survival skills for hunter-gatherers, are now being applied to data sets rather than animal behaviours and the outcome is insight into behaviour.

2.5 Literature summary

BD describes the large quantity of data generated and stored by modern technologies which have forced businesses to consider how they exploit the resultant information flows. Collecting data from multiple channels has never been easier. The idea that data shadows and information trails of people, machines, commodities and even nature can reveal secrets because the power and prowess now exists to uncover them is alluring. The hype says that the more data gathered, the better the decision making. Reality is less straightforward. Usability is critical. Data need to be converted in a robust and reliable way to be translated into knowledge to be applied. Agility and flexibility in data collection is good, but it is

necessary to connect and correlate relationships, hierarchies and multiple data linkages, otherwise the data remain meaningless. Meaningless data are the digital information equivalent of a rubbish heap, as shown by the GFT example.

In an ideal world, BD would help organisations set new levels of performance and new agendas for business. This research explores how advances are being made that make this a possibility and considers the challenges that remain.

3. Methodology

The research objective was to understand if organisations use BD and if its benefits match its widely presented potential. Given that this is a relatively new research area and the research largely exploratory, to achieve the objectives a two-phase multiple case study approach was adopted (McCutcheon and Meridith, 1993). In the first phase secondary data from BD vendors informed the development of a theoretical framework categorising current BD usage. Phase 2 explored two organisations that have sought to exploit BD to evaluate whether the framework is supported by empirical data.

3.1 Phase 1

Products that are available influence how organisations are likely to exploit BD. An exploration of the types of application that BD vendors offer brings insights into the ways that BD is being used. A purposeful sampling strategy was adopted because it was sympathetic to the objectives and the exploratory nature of the research. An earlier study had already identified the top 10 vendors of BD solutions (Kelly, 2014). This was validated against other sources of data (industry and market reports, trade journals and media coverage) and judged to remain valid for the purposes of this research and was therefore the sample used.

Individual case studies produced by the vendor and made available on their website that related only to applications of BD were examined. To mitigate against the risk of obsessive positivity, only case studies with a named client were used, based on the premise that the client would not allow their reputation to be damaged by exaggerated vendor claims. A paragraph describing each case was created.

Cases relating solely to the implementation of technology infrastructure were rejected, as this is independent of the application. Three of the vendors did not provide case studies or provided them in an anonymous form, restricting the vendor sample to seven. Within this sample, some vendors provided numerous case studies and others few. A total of 253 case studies were identified across a wide range of industrial sectors. There was deliberately no attempt to establish in-case validity or consistency through triangulation. Rather, this phase of the research sought to identify the range of product offerings, not to classify the vendors' products. This may have led to an unrepresentative data set, but establishing the range of applications was the objective at this stage.

Eisenhardt (2002) indicates that between four and ten case studies is an ideal number to develop a theory, so numerically the validity of the approach adopted is high. Grounded theory is used to build a framework of BD usage based on the client track record of the leading BD vendors, in keeping with Glaser and Strauss (1968).

With the number of usable case studies amongst the vendors varying widely, it was necessary to mitigate bias towards any particular vendor based on proclivity to publish cases. Client track record is primarily a mechanism for marketing the services of a professional services firm or technology vendor. The case studies that comprise their track record are therefore universally positive and likely to be skewed towards applications that are more profitable. Consequently, a limitation of the methodology is that unprofitable implementations of BD may be missing from any analysis undertaken. This is an area for further research.

3.2 Phase 2

The second phase of the research examined the experiences of two UK retail organisations as they embarked on performance improvement projects using BD. In this instance the sampling was opportunistic. Readily available data and direct access rather than reliance on self-reporting were more important factors than an alternative sampling strategy.

The case studies, their drivers, method and outcome are outlined separately. They are compared with the general literature on BD, what it is and what it is supposed to do (Wieland *et al.*, 2014; Fosso Wamba *et al.*, 2015; Haas *et al.*, 2014). The cases are applied to the vendor analysis framework developed from the earlier phase. The aim was to ascertain how organisations currently leverage BD and if there is anything to differentiate this from standard data analysis techniques.

Aligning the two phases, the research sought to compare the hype with the reality in an operational setting. Thus the operational implications of BD usage for sustainable competitiveness are highlighted whilst considering the social, technological and human consequences, in line with Kranzberg (1986). As such the research complements the work undertaken by Dubey *et al.* (2016). Both adopt a two-phase, mixed-methods, sequential approach of theory-building, followed by the evaluation of primary data to provide a degree of theory validation.

4. Framework development: the applications of BD

Content analysis was undertaken for each of the descriptive paragraphs of the case studies, and a list of categories developed. Each case study was categorised with one or more types. Some had several threads, so were treated as more than one content item. Whilst some categories were far more prevalent than others in the published cases, all applications, even if there was only a single example, were incorporated into the analysis. Hence, by controlling for the popularity of an application, the research covered all applications equitably. A set of aggregate classifications was developed and again each paragraph was classified.

The categorisation identified a number of application areas which resulted in the development of a framework on two axes: scale and time horizon, shown in Table I.

The scale axis considered whether the case study provided BD services that focussed on macro applications, e.g. analysing flows within a supply chain, or micro applications, e.g. providing recommendations to a single customer. The time horizon scale looked at whether the focus of the data were regarding the past, for example, historical data analysis to demonstrate compliance, the present, for example, real-time fraud detection, or modelling the future, for example, new service proposition design. Table II shows the full description for all the categories. The framework was then validated against the stated service offerings of the vendors to ensure that all of their offerings could be classified by within the typology. This was done by reviewing the section of vendor websites relating to services provided and plotting each service back onto the framework.

Table I.
Typology of BD
application

Scale	Micro (1)	Macro (2)
<i>Time horizon</i>		
Future	Prediction	Service design strategic support Optimisation
Present	Personalisation Detection	
Retrospective	Troubleshooting	Compliance

Table II.
Description of
typological categories

Category	Description
Prediction	Using BD to predict what products an individual customer would want in future, for example, the promotion of recommended purchases
Service design	Using the data feeds from existing products and services to inform the design of new product or service propositions
Strategic support	Using BD to inform strategic decisions
Personalisation	Using BD to personalise the experience of an individual customer, for example, the provision of information that they were likely to want
Detection	The use of BD to identify issues in real-time, for example, fraud detection
Optimisation	The use of BD to optimise performance in a process, for example, through its use in supply chain planning
Troubleshooting	Use of BD to identify the causes of problems that have occurred
Compliance	Use of BD to demonstrate compliance to regulations within an organisation. This includes using BD to analyse complex sources of data such as recordings of telephone conversations

This typology provides a framework for the classification of the application of BD in operations management and the wider field of management. The next step was to test its validity by examining two case studies.

5. The case studies built on primary data

5.1 Case study 1 – online white goods retailer

Background. This organisation wanted to deliver tailored experiences for new online customers when they visited the company's website. The absence of physical clues about a customer such as gender, clothing and ethnicity that are used in face-to-face encounters meant that the statistically significant real-time sales discrimination based on appearance was not possible (Wise, 1974). An employee was musing one day and wondered if, in the absence of physical clues, the technology the customer used could help the company understand the customer better. The research question was:

RQ1. Can historical access data deliver more tailored experiences to new customers when they visit the website?

Data collection. Search advertising is the primary tool directing potential customers to the website and is complemented by display advertising where defined criteria about context and previous browsing behaviour are met. Conversion rate optimisation (CRO) is used to increase the number of website visitors booking a design visit and thereby moving more successfully through the company's sales funnel. CRO positively influences cost-per-lead and permits analysis of the technology in the form of operating system (OS) and browser customers use to access the website. OS and browser data combined was collected for a 12-month period (2014). In total, 1,621,262 website clickstream data were pulled into the company's CRM system at the point of design visit enquiry, and analysed.

Data analysis. Using Google analytics to investigate internet traffic data it became apparent that the research question could be more precisely defined:

RQ2. Can customer technology use for website navigation predict purchase type and spend?

RQ3. Can this information be used to deliver a more tailored website experience?

Preliminary analysis showed that Microsoft Explorer (IE) and Firefox usage declined by 40 per cent in 12 months while all other browser traffic increased. Devices were identified as the most important influencer because most users stay with the default browser provided

(Browser-update.org, 2015). As use of smartphones and tablets increased, so did access to the company website using Safari and Google Chrome, with corresponding decreases in IE and Firefox OS.

Three separate hypotheses were developed to test the correlation between purchasing choice and variables associated with web use that could be captured using Google analytics:

- H1. Product style preference and device are associated.
- H2. Product style preference and OS are associated.
- H3. Product style preference and browser are associated.

The clickstream data were segmented into three technology categories and number of visits to the top 4 styles of the two product categories the company offered. Pearson's χ^2 test of association was selected to test the hypotheses because of its versatility and ability to deal with categorical data (Hair *et al.*, 2007). Statistical significance was tested to 95 per cent, judged by the organisation to be an appropriate level of confidence to use in the tests (Table III).

Thus it was possible to accept all three hypotheses.

Because of the interdependency between the OS and browser, a fourth hypothesis was developed:

- H4. Product style preference and [OS + browser] are associated.

A total of 21 different combinations of OS and browser were identified from the data available. Again using Pearson's χ^2 test and the same significance level as before, a p value of 6.1098×10^{-195} indicated that this hypothesis could also be accepted. Similarly the critical value of 31.4104 was significantly below the test statistic (χ^2) of 980.3270.

<i>H1 There is association between product preference and the device used to browse</i>	
<i>H₀</i>	Product preference and device are not associated
<i>H1</i>	Product preference and device are associated
Significance level	$\alpha = 0.05$
Pearson's χ^2 test	$(\chi^2) = 509.2441896$
Test statistics	$P = 2.6244 \times 10^{-111}$
Evaluation	Critical test = 5.991464547
	$(\chi^2) > \text{Critical test } p > \alpha$
	From both tests reject H_0 and accept $H1$
<i>H2 There is an association between product preference and the operating system used whilst browsing</i>	
<i>H₀</i>	Product preference and operating system are not associated
<i>H1</i>	Product preference and operating system are associated
Significance level	$\alpha = 0.05$
Pearson's χ^2 test	$(\chi^2) = 839.9343174$
Test statistics	$P = 4.4630 \times 10^{-177}$
Evaluation	Critical test 14.06714045
	$(\chi^2) > \text{Critical test } p > \alpha$
	From both tests reject H_0 and accept $H1$
<i>H3 There is an association between product preference and the browser used</i>	
<i>H₀</i>	Product preference and browser are not associated
<i>H1</i>	Product preference and browser are associated
Significance Level	$\alpha = 0.05$
Pearson's χ^2 test	$(\chi^2) = 685.4852838$
Test statistics	$P = 6.7527 \times 10^{-146}$
Evaluation	Critical test 11.07049769
	$(\chi^2) > \text{Critical test } p > \alpha$
	From both tests reject H_0 and accept $H1$

Table III.
Case study 1:
hypotheses testing

Findings. The result from *H4* definitively answers *RQ2* and confirms that assessing customer preference based on technology is viable. Combining OS and browser variables provided a rich set of data segments to use in tailoring the customer experience. Mac OS and Google Chrome suggested a preference for a modern, expensive products whilst IE meant more traditional style and careful spend. *RQ3* was also answered from test 4. The strong association between browser/OS combinations and product style preferences enabled the business to identify three customer categories based on the technologies used for browsing:

- (1) more interested in traditional;
- (2) more interested in modern; and
- (3) no bias.

The website was redeveloped to act upon insights produced from the statistical analysis of browsing behaviour data. By collecting information about the OS and browser combination during the initial online contact and applying the knowledge of likely customer preferences, customers could be direct towards specific webpages tailored to meet their likely preferences. Customers benefited by quicker navigation to the products they were likely to be interested in, with a corresponding improvement in the “hits: design visit” ratio.

Considering the case carefully, it is clear that the analysis undertaken was relatively straightforward, but the improvement in sales would not have been possible without the right question being asked and an association being found between hitherto unmade combinations.

When assessed according to the framework presented in Section 4, this retailer’s approach falls into the category of personalisation.

5.2 Case study 2 – multi-channel retailer

This company had a vision: to create a competitive edge by being the most trusted UK provider in the sector. This retailer has 19 branches nationwide and several brand names linked to it, with three divisions: inbound, re-sales and retail. The inbound creates the majority of activity, provides the stock and accounts for 10 per cent of revenue. Re-sales represent 70 per cent of the revenue. Retail reaches the consumer directly and accounts for the remaining 20 per cent of revenue.

Many years’ data from multiple outlets tracking all sales operations existed and the company wanted to set up a “scientific platform” to analyse it. In 2014 the data were interrogated differently for the first time. The first problem to arise was variety – inconsistent, incomplete and inaccurate data capture across outlets and delivery channels. Despite being internally generated, data had to be cleaned because of non-standardised data capture processes. There were two product ranges, value and deluxe. Historically the deluxe range had been “enhanced” by providing a six-month guarantee. They were surprised to find they lost money on the deluxe and made three times more, per value item, even though it sold more cheaply and it did not have a guarantee. An unexpected pattern in sales was also found: certain models sold for higher prices in certain locations, yet their sales policy was to sell at the location nearest to the previous owner’s registered address, regardless of sales value location. Stock days were found to be irrelevant.

These findings meant that the management team reconsidered the positioning of the product range and standardised a number of aspects of their operation hitherto non-standardised.

Data collection. The operation generates masses of data in a constant stream across all its business units. Data come from “in use”, “off use”, retail, wholesale, industry regulators and each individual product item. Total sales are approximately 500,000 units annually. Wanting to create a competitive edge by becoming the most trusted UK provider in its

sector the company set up a project to provide a scientific platform to analyse the data collected from the day-to-day business operations. The research questions were brainstormed by the sales team, led by the heads of sales and sales effectiveness. The aim was to understand the impact of product variety on workload and profitability. The research questions which emerged were:

RQ4. Are products with guarantees more profitable?

RQ5. What links are there between wholesale and retail?

RQ6. What is the link between stock days and product profile?

Hypotheses were developed to run the DM verification paradigm as well as enable descriptive analysis:

H5. Product with guarantees are more profitable than those without.

H6. Deluxe products are more profitable than standard.

The re-sales division had the largest data set. An initial exploration revealed enough data sets to answer the research questions.

Data preparation. A total of 48 variables were found within the sales data. To enable subsequent analysis the variables were defined into: numeric, categorical (string), Boolean (yes/no) and specific codification attributes. Data cleaning had to be carried out due to three main inputting problems – inconsistency (e.g. BLK and BLACK), character transposition and missing values.

Data analysis. The company adopted the cross industry standard process for data mining approach as the most viable to mine existing data because it is accepted as the “gold standard” in the DM domain (Rennolls and Al-Shawabkeh, 2008; Pechenizkiy *et al.*, 2008).

Using this approach it was found that non-guaranteed products’ mean contribution to profitability was higher than those with guarantees, therefore *H5* was rejected. Total revenue from deluxe products was found to be 61.5 per cent, therefore *H6* was also rejected. Further exploration of the deluxe category established there was product differentiation within-category. The DM findings identified the most profitable sales.

Findings. The insight that inherently enhanced products were more profitable than those enhanced by guarantee led the management team to reconsider the market positioning of the product range and the input the company provided pre-sales. The descriptive analysis carried out to answer *RQ5* highlighted the most searched products nationally, by location, which can help decision making regarding targeted pricing and promotion decisions. The *RQ6* analysis ascertained there was no particular link between sales and stock days.

The findings from this analysis of all-company data provide a framework for decision making regarding both strategic and tactical aspects of the business. In addition, it has led to standardisation of data capture to reduce future inconsistency. The company has also recognised that having developed an analysis platform there are many other aspects it can explore operationally. Having established an industry overview, further mining of the data can help it enhance its market leader position. In terms of the framework, this falls into the category of strategic support, and containing aspects of service design.

6. Discussion

Whilst there are limitations within both case studies, they illustrate that there is promise in using BD to uncover previously unavailable insights for sustainable performance.

Both case studies conformed to the Gartner (2014) online definition of BD: “data assets that require innovative forms of information processing for enhanced insight and decision-making”. Case 1 highlighted the importance of knowing your product and what

exactly you wanted to know about your customer and the value of asking interesting and insightful questions. Case 2 highlighted the difficulties encountered with inconsistent data and the importance of reliability and robustness, and showed how to maximise service levels when you use the complete knowledge of long-term process of new sales and used sales over a period of years. These case-specific points relate to the respective organisations, although they can generally be seen to be in keeping with the OECD (2013) suggestion and confirm the importance of developing the correct strategy to advantageously exploit the data (Opresnik and Taisch, 2015). They are also problems consistent with the core concerns of operations management which scholars have studied continuously for many years (Ackoff, 1989; Kranzberg, 1986; Bell, 1999; Saaty, 2008). It is this aspect which has led to observations that BD is another step in data-driven improvement, following in the footsteps of SPC and then Six Sigma (Madsen and Stenheim, 2013; Näslund, 2008).

The analysis of the competitive landscape relating to BD vendors enables a more granular definition of “insight and decision-making”. Each of the two case studies demonstrated different elements of the BD framework in use. The framework could be used to guide the business in Case 1 to extend their use of personalisation into the area of prediction of product bundles that will be desired by customers, and thence to improved service design and responsiveness. Equally the business in Case 2 was described as beginning to use its BD-driven insights to redesign its portfolio of services and to apply these in a targeted way to product categories, an extension of BD use predicted by the framework.

Case 1 wanted to enhance the customer experience. Case 2 wanted to become the most trusted UK provider in its market. Both companies used internal data about committed customers to create competitive advantage. Unlike GFT, where individuals other than those suffering flu would make searches about the topic, in this case the stream of data was a complete picture of the companies’ customers and their behaviour. The use of BD techniques meant that it was unnecessary to draw samples of behaviour and try to use this to predict individual responses. By using BD techniques both companies were able to maximise their management decisions, based on a complete picture of the behaviour of their total customer base. This is a distinctly better method of analysis than simply looking for trends in the whole population and trying to relate it to their business, as was being done with GFT. It is also a true application of BD, whereas the GFT case, despite collating social media and other externally generated data were actually using a population sample. Sampling and hypothesis-building is traditional data analysis, not BD, despite the constant stream of largely unstructured data created from multiple sources and in a range of formats which causes the BD analytics challenge (Bisel *et al.*, 2014; Chen *et al.*, 2012; Tinati *et al.*, 2014).

Case 2’s data, whilst not quite unstructured, lacked a standardised company format. The same products were differently captured on the system by each branch, creating a cleaning project first, then a project to systematically identify capture data throughout the organisation, regardless of product or channel.

The key aspect for both cases is the importance of initial inspiration by individual employees. Neither case would have happened had a member of staff not wondered about a particular problem, just as with Wald in 1943. Technology enables the speedier realisation of an answer but it cannot show the direction to choose for exploration. So, far from being helpful, BD, as stated in the literature review, has enormous potential for misinterpretation and misdirection, of which numerous examples exist (Lazer *et al.*, 2014; Ilic *et al.*, 2010; Hazen *et al.*, 2014; Duhigg, 2012). Whilst the framework developed from this research can help, it is specific people skills which are required to outline the initial enquiry. There is a clear need, as shown by the case studies, to have employees with a level of curiosity which can lead to new streams of exploration. Relying on algorithms and system tools is unlikely to present new insights. The key challenge therefore is to identify the new skills people need in order to be able to maximise the potential BD offers and underlines the relevance of the GCHQ development (Philipson, 2014).

These two empirical studies are the first step in validating the BD framework, which can help organisations in two key ways. First, it can help guide the nature of the questions required based on the outcome sought. Second, this helps shape an understanding of the people skills required for both interpreting data sets and developing relevant, impactful insight. The white goods retailer highlighted the importance of knowing your product and what exactly you wanted to know about your customer. The mixed retailer data analysis highlighted the difficulties encountered with inconsistent data and the importance of reliability and robustness. More examples and tests are needed.

7. Conclusion

This paper has explored the challenges and opportunities of BD. The research shows that despite the hype BD solutions are being used to manage and improve operational processes. Irrespective of the longevity of the term and the novelty of the techniques, the data available to organisations will continue to grow and technology will be increasingly utilised to harness the value locked away in that data. The challenge to harnessing the full power of the technology is in formulating questions and new perspectives from which to examine problems or perceived problems and the answers to these problems. This then enables organisations to exploit and leverage the power of the technology to create sustainable competitive advantage. The framework provides a conceptual model to target desired outcomes and the questions to get there.

The collection, processing and utilisation of data to inform decision making and optimise processes are defining features of operations management and operations research. Operations are already being controlled and automated using advanced technologies and the availability of more data and sophisticated analytic tools will be exploited in the future. Yet limitless data do not guarantee better operational performance. Operations managers and researchers need to learn how to use new tools and develop the cognitive skills necessary to generate the knowledge and wisdom they will require to manage operations within an increasingly connected, complex and rapidly changing world.

BD is identified in three main areas: administrative (from public bodies), social media and private sector data. A number of problems emerge, which are also opportunities for further research. There are access and reliability problems. Understanding the utility of the data as research resources is paramount. Given the variety of sources, data reuse, linkage and sharing may be difficult. How to overcome these issues, and uncover if it is possible to do so reliably, is a fruitful avenue for future researchers.

There are problems with how people think, react and respond to things at any given time. For instance the way speed is seen affects how data are perceived. Speed may mean better, more informed decisions can be made by individuals for themselves and for their organisations. Or it may mean greater confusion due to rapidly changing scenarios having to be grappled with. Operations management capacity issues now need to also incorporate how much data people can cope with to optimise what is done and how it is done. Behavioural operations have a key role to play in researching whether behaviour is influenced by the data presented or whether behaviour affects the data. Changes in how customers are approached and dealt with also need to be considered. Just as organisations are impacted by BD, so are customers. They have more information at their fingertips than ever before. Speed becomes a factor for them too, as does knowledge. Organisations need to understand what customers do with the knowledge they have, how they exercise their decision-making choices and if/how organisations can influence that.

Staff requirements also need to be rethought, and research into this is paramount. There needs to be a step change in competences for some of the tasks expected to be done as a matter of course once the world of BD is embraced. New data science skill-sets are required, because current practices, without insightful and curious employees, mean more data are

being churned more rapidly but not necessarily for greater advantage. Indeed, the GFT example shows how it can impede true knowledge by being inherently flawed. The key challenge therefore is to identify the new skills people need in order to be able to maximise the potential BD offers, as GCHQ have already done, although much more work is needed.

Finally, data curation needs have to be rethought. As yet there is no evidence how this is going to be tackled in a sustainable way, or indeed if it will be.

The framework proposed in this study offers a potentially useful starting point for understanding where and how organisations can leverage business value from BD. The case studies demonstrate two of the categories in use. Further research is required to fully validate the remaining categories within the framework and to ascertain its utility. Future organisational success, if not survival, may well be predicated on insightful application of systems-generated, people-enhanced information.

References

- Ackoff, R. (1989), "From data to wisdom", *Journal of Applied Systems Analysis*, Vol. 16, pp. 3-9.
- Barton, D. and Court, D. (2012), "Making advanced analytics work for you", *Harvard Business Review*, Vol. 90 No. 10, pp. 78-83.
- Bell, P. (2013), "Creating competitive advantage using big data", *Ivey Business Journal*, Vol. 77 No. 3, pp. 4-8.
- Bell, P.C. (1999), *Management Science/Operations Research: A Strategic Perspective*, South-Western College Publishing, Boston, MA.
- Bisel, R.S., Barge, J.K., Dougherty, D.S., Lucas, K. and Tracy, S.J. (2014), "A round-table discussion of 'big' data in qualitative organizational communication research", *Management Communication Quarterly*, Vol. 28 No. 4, pp. 625-649.
- Bollen, J., Mao, H. and Zeng, X. (2011), "Twitter mood predicts the stock market", *Journal of Computational Science*, Vol. 2 No. 1, pp. 1-8.
- Brown, B., Chui, M. and Manyika, J. (2011), "Are you ready for the era of 'big data'?", *McKinsey Quarterly*, Vol. 4, pp. 24-35.
- Browser-update.org (2015), "Homepage statement", available at: www.linkedin.com/company/6202954?trk=vsrp_companies_cluster_name&trkInfo=VSRPsearchId%3A32485231427721498574%2CVSRPtargetId%3A6202954%2CVSRPcmpt%3Acompanies_cluster (accessed 30 March 2015).
- Bughin, J., Livingston, J. and Marwaha, S. (2011), "Seizing the potential of 'big data'", *McKinsey Quarterly*, Vol. 4, pp. 103-109.
- Carvajal, R. (1992), "Operations research (or), management science (Ms), systems science, and Russell Ackoff: the development of two paradigms", *Systems Practice*, Vol. 5 No. 3, pp. 291-318.
- Chae, B. (2015), "Insights from hashtag #supplychain and Twitter analytics: considering Twitter and Twitter data for supply chain practice and research", *International Journal of Production Economics*, Vol. 165, pp. 247-259.
- Chen, H., Chiang, R.H.L. and Storey, V.C. (2012), "Business intelligence and analytics: from big data to big impact", *MIS Quarterly*, Vol. 36 No. 4, pp. 1165-1188.
- Chongwatpol, J. and Chan, H.K. (2015), "Prognostic analysis of defects in manufacturing", *Industrial Management & Data Systems*, Vol. 115 No. 1, pp. 64-87.
- Cox, M. and Ellsworth, D. (1997), "Application-controlled demand paging for out-of-core visualization", *Proceedings of the 8th Conference on Visualization '97*, IEEE Computer Society Press, Phoenix, AZ, pp. 235-243.
- Croxall, B. (2014), "The Red Herring of Big Data, Brian Croxall web post", available at: www.briancroxall.net/2014/03/07/the-red-herring-of-big-data/ (accessed 30 March 2015).
- Demirkan, H. and Delen, D. (2013), "Leveraging the capabilities of service-oriented decision support systems: putting analytics and big data in cloud", *Decision Support Systems*, Vol. 55 No. 1, pp. 412-421.

- Diebold, F.X. (2012), "A personal perspective on the origin(S) and development of 'Big Data': the phenomenon, the term, and and the discipline", PIER Working Paper Nos 13-003, second version, Department of Economics, University of Pennsylvania, PA, 26 November.
- Dubey, R., Gunasekaran, A., Childe, S.J., Wamba, S.F. and Papadopoulos, T. (2016), "The impact of big data on world-class sustainable manufacturing", *The International Journal of Advanced Manufacturing Technology*, Vol. 84 No. 1, pp. 631-645.
- Duhigg, C. (2012), *How Companies Learn Your Secrets*, NY Times, available at: www.nytimes.com/2012/02/19/magazine/shopping-habits.html?pagewanted=1&_r=2&hp& (accessed 12 December 2015).
- Eisenhardt, K.M. (2002), "Building theories from case study research", in Huberman, A.M. and Miles, M.B. (Eds), *The Qualitative Researcher's Companion*, Sage Publications, Thousand Oaks, CA and London, pp. 5-35.
- Esener, S.C., Kryder, M.H., Doyle, W.D., Keshner, M., Mansuripur, M. and Thompson, D.A. (1999), *The Future of Data Storage Technologies*, PB99-144214, International Technology Research Institute, Baltimore, MD.
- Fawcett, S.E. and Waller, M.A. (2014), "Supply chain game changers – mega, nano, and virtual trends – and forces that impede supply chain design (i.e. building a winning team)", *Journal of Business Logistics*, Vol. 35 No. 3, pp. 157-164.
- Fisher, D., DeLine, R., Czerwinski, M. and Drucker, S. (2012), "Interactions with big data analytics", *Interactions*, Vol. 19 No. 3, pp. 50-59.
- Fosso Wamba, S., Akter, S., Edwards, A., Chopin, G. and Gnanzou, D. (2015), "How 'big data' can make big impact: findings from a systematic review and a longitudinal case study", *International Journal of Production Economics*, Vol. 165, July, pp. 234-246.
- Gandomi, A. and Haider, M. (2015), "Beyond the hype: big data concepts, methods, and analytics", *International Journal of Information Management*, Vol. 35 No. 2, pp. 137-144.
- Gartner (2013), "Emerging technologies hype cycle. Gartner research and advisory", available at: www.gartner.com/newsroom/id/2575515 (accessed 14 September 2015).
- Gartner (2014), "Gartner says annual smartphone sales surpassed sales of feature phones for the first time in 2013", Gartner, available at: www.gartner.com/newsroom/id/2665715 (accessed 13 February 2014).
- Gartner (2015), "Emerging technologies hype cycle. Gartner research and advisory", available at: www.gartner.com/newsroom/id/3114217 (accessed 14 September).
- George, G., Haas, M.R. and Pentland, A. (2014), "Big data and management", *Academy of Management Journal*, Vol. 57 No. 2, pp. 321-326.
- Glaser, B.G. and Strauss, A.L. (1968), *The Discovery of Grounded Theory: Strategies for Qualitative Research*, Weidenfeld & Nicolson, London.
- Haas, M.R., Criscuolo, P. and George, G. (2014), "Which problems to solve? Attention allocation and online knowledge sharing in organizations", *Academy of Management Journal*, Vol. 58 No. 3, pp. 680-711.
- Hair, J.F.H., Money, A.H., Samouel, P. and Page, M. (2007), *Research Methods for Business*, Wiley, Chichester.
- Hazen, B.T., Boone, C.A., Ezell, J.D. and Jones-Farmer, L.A. (2014), "Data quality for data science, predictive analytics, and big data in supply chain management: an introduction to the problem and suggestions for research and applications", *International Journal of Production Economics*, Vol. 154, pp. 72-80.
- Housley, W., Procter, R., Edwards, A., Burnap, P., Williams, M., Sloan, L., Rana, O., Morgan, J., Voss, A. and Greenhill, A. (2014), "Big and broad social data and the sociological imagination: a collaborative response", *Big Data & Society*, Vol. 1 No. 2, pp. 1-15.
- Hruska, J. (2013), "Intel's former chief architect: Moore's law will be dead within a decade", ExtremeTech, available at: www.extremetech.com/computing/165331-intels-former-chief-architect-moores-law-will-be-dead-within-a-decade (accessed 30 March 2015).

- Huang, Y.-Y. and Handfield, R.B. (2015), "Measuring the benefits of ERP on supply management maturity model: a 'big data' method", *International Journal of Operations & Production Management*, Vol. 35 No. 1, pp. 2-25.
- IDC (2012), "The digital universe in 2020: big data, bigger digital shadows, and biggest growth in the far east", IDC, available at: www.emc.com/collateral/analyst-reports/idc-the-digital-universe-in-2020.pdf (accessed 13 February 2014).
- Ilic, A., Grössbauer, A., Michahelles, F. and Fleisch, E. (2010), "Understanding data volume problems of RFID-enabled supply chains", *Business Process Management Journal*, Vol. 16 No. 6, pp. 904-916.
- Jifa, G. and Lingling, Z. (2014), "Data, Dikw, big data and data science", *Procedia Computer Science*, Vol. 31, pp. 814-821.
- Kelly, J. (2014), "Big data vendor revenue and market forecast 2013-2017. Wikibon", available at: http://wikibon.org/wiki/v/Big_Data_Vendor_Revenue_and_Market_Forecast_2013-2017 (accessed 19 March 2015).
- Kranzberg, M. (1986), "Technology and history: 'Kranzberg's Laws'", *Technology and Culture*, Vol. 27 No. 3, pp. 544-560.
- Laney, D. (2001), "3D data management: controlling data volume, velocity and variety", Application Delivery Strategies, 4, available at: <http://blogs.gartner.com/doug-laney/files/2012/01/ad949-3D-Data-Management-Controlling-Data-Volume-Velocity-and-Variety.pdf> (accessed 19 March).
- Lazer, D., Kennedy, R., King, G. and Vespignani, A. (2014), "The parable of Google flu: traps in big data analysis", *Science*, Vol. 343 No. 6176, pp. 1203-1205.
- Lee, H. and Özer, Ö. (2007), "Unlocking the value of RFID", *Production and Operations Management*, Vol. 16 No. 1, pp. 40-64.
- Li, X., Song, J. and Huang, B. (2016), "A scientific workflow management system architecture and its scheduling based on cloud service platform for manufacturing big data analytics", *The International Journal of Advanced Manufacturing Technology*, Vol. 84 No. 1, pp. 119-131.
- Liew, P. (2007), "Understanding data, information, knowledge and their inter-relationships", *Journal of Knowledge Management Practice*, Vol. 8 No. 2, available at: www.tlinc.com/jkmpv8n207.htm
- Lucke, D., Constantinescu, C. and Westkämper, E. (2008), "Smart factory – a step towards the next generation of manufacturing", in Mitsuishi, M., Ueda, K. and Kimura, F. (Eds), *Manufacturing Systems and Technologies for the New Frontier*, Springer, London, pp. 115-118.
- Lury, C. (2013), "Focus: big data, little questions?", *Discover Society*, Vol. 1, pp. 1-6.
- McAfee, A. and Brynjolfsson, E. (2012), "Big data: the management revolution (cover story)", *Harvard Business Review*, Vol. 90 No. 10, pp. 60-68.
- McCutcheon, D.M. and Meredith, J.R. (1993), "Conducting case study research in operations management", *Journal of Operations Management*, Vol. 11 No. 3, pp. 239-256.
- Madsen, D.Ø. and Stenheim, T. (2013), "Doing research on 'management fashions': methodological challenges and opportunities", *Problems and Perspectives in Management*, Vol. 11 No. 4, pp. 68-76.
- Manyika, J., Chui, M., Brown, B., Bughin, J., Dobbs, R., Roxburgh, C. and Hung Byers, A. (2011), *Big Data: The Next Frontier for Innovation, Competition and Productivity*, McKinsey Global Institute, available at: www.mckinsey.com/business-functions/digital-mckinsey/our-insights/big-data-the-next-frontier-for-innovation
- Meyer, G.G., Buijs, P., Szirbik, N.B. and Wortmann, J.C. (2014), "Intelligent products for enhancing the utilization of tracking technology in transportation", *International Journal of Operations & Production Management*, Vol. 34 No. 4, pp. 422-446.
- Miller, S. (2014), "Collaborative approaches needed to close the big data skills gap", *Journal of Organization Design*, Vol. 3 No. 1, pp. 26-30.
- Moore, G.E. (1965), "Cramming more components onto integrated circuits", *Electronics*, Vol. 38 No. 8, pp. 114-117.

- Näslund, D. (2008), "Lean, Six Sigma and lean sigma: fads or real process improvement methods?", *Business Process Management Journal*, Vol. 14 No. 3, pp. 269-287.
- OECD (2013), "Exploring data-driven innovation as a new source of growth: mapping the policy issues raised by 'big data'", OECD Digital Economy Papers, No. 222, OECD Publishing.
- O'Leary, D.E. (2013), "'Big data', the 'internet of things' and the 'internet of signs'", *Intelligent Systems in Accounting, Finance and Management*, Vol. 20 No. 1, pp. 53-65.
- Opresnik, D. and Taisch, M. (2015), "The value of big data in servitization", *International Journal of Production Economics*, Vol. 165, July, pp. 174-184.
- Pechenizkiy, M., Puuronen, S. and Tsymbal, A. (2008), "Towards more relevance-oriented data mining research", *Intelligent Data Analysis*, Vol. 12 No. 2, pp. 237-249.
- Philpison, A. (2014), "GCHQ employs more than 100 dyslexic and dyspraxic spies", available at: www.telegraph.co.uk/education/educationnews/11111584/GCHQ-employs-more-than-100-dyslexic-and-dyspraxic-spies.html (accessed 19 January 2015).
- Piccoli, G. and Pigni, F. (2013), "Harvesting external data: the potential of digital data streams", *MIS Quarterly Executive*, Vol. 12 No. 1, pp. 53-64.
- Porter, M.E. and Heppelmann, J.E. (2014), "How smart, connected products are transforming competition", *Harvard Business Review*, Vol. 92 No. 11, pp. 64-88.
- Rennolls, K. and Al-Shawabkeh, A. (2008), "Formal structures for data mining, knowledge discovery and communication in a knowledge management environment", *Intelligent Data Analysis*, Vol. 12 No. 2, pp. 147-163.
- Rowley, J. (2007), "The wisdom hierarchy: representations of the DIKW hierarchy", *Journal of Information Science*, Vol. 33 No. 2, pp. 163-180.
- Saaty, T.L. (2008), "Decision making with the analytic hierarchy process", *International Journal of Services Sciences*, Vol. 1 No. 1, pp. 83-98.
- Sætre, A.S., Sørnes, J.-O., Browning, L.D. and Stephens, K.K. (2003), "Organizational members' enactment of organizational environments and media use: a study of ICT practices in Norway and the United States", *Informing Science and Information Technology Education Joint Conference, Pori*.
- Segerberg, A. and Bennett, W.L. (2011), "Social media and the organization of collective action: using Twitter to explore the ecologies of two climate change protests", *The Communication Review*, Vol. 14 No. 3, pp. 197-215.
- Shedroff, N. (1994), "Information interaction design a unified field theory of design", in Jacobson, R. (Ed.), *Information Design*, MIT Press, p. 373.
- Tinati, R., Halford, S. and Carr, L. (2014), "The promise of big data: new methods for sociological analysis", *Sociology*, Vol. 48 No. 4, pp. 663-681.
- Wieland, A., Handfield, R.B. and Durach, C.F. (2014), "Mapping the landscape of emerging research topics in supply chain management", *Proceedings of 21st EurOMA Conference, European Operations Management Association, Palermo and Sicily*, 9pp.
- Wise, G.L. (1974), "Differential pricing and treatment by new-car salesmen: the effect of the prospect's race, sex, and dress", *Journal of Business*, Vol. 47 No. 2, pp. 218-230.
- Zelbst, P.J., Green, K.W., Sower, V.E. and Reyes, P.M. (2012), "Impact of RFID on manufacturing effectiveness and efficiency", *International Journal of Operations & Production Management*, Vol. 32 No. 3, pp. 329-350.
- Zhong, R.Y., Lan, S., Xu, C., Dai, Q. and Huang, G.Q. (2016), "Visualization of RFID-enabled shopfloor logistics big data in cloud manufacturing", *The International Journal of Advanced Manufacturing Technology*, Vol. 84 No. 1, pp. 5-16.

Further reading

- Wald, A. (1980), *A Reprint of A Method of Estimating Plane Vulnerability Based on Damage of Survivors*, CRC 432, Center for Naval Analyses: Operations Evaluation Group, Alexandria, VA.

About the authors

Olga Matthias earned her DBA from the Bradford University and is currently a Senior Lecturer at the School of Management. Her research interests include the co-dependency between relationship and service quality in management consultancy, knowledge creation and relationships in B2B professional services. Olga Matthias is the corresponding author and can be contacted at: o.matthias1@bradford.ac.uk

Ian Fouweather is an Associate Lecturer at the Faculty of Management and Law. His interests include change management, sensemaking, eCommerce, management education, The Philosophy of Research and Linguistic Philosophies.

Ian Gregory is an Associate Lecturer in Operations Management and Information Systems at the University of Bradford. His research interests include software development, knowledge transfer and the use of IT systems to improve operational performance.

Andy Vernon is a member of PA Consulting Group's Management team. He is a Trusted Advisor to many clients, with 30 years' experience of helping organisations use digital technologies to transform their services, operations and culture. Andy combines deep change management, IT and business strategy expertise to make business transformations a reality.

This article has been cited by:

1. Fosso Wamba Samuel Samuel Fosso Wamba fossowam@gmail.com Samuel Fosso Wamba, PhD, HDR, is a Full Professor at the Toulouse Business School, France in September 2016. Prior, he held several positions including Associate Professor at the NEOMA Business School, France and Senior Lecturer at the School of Information Systems & Technology (SISAT), University of Wollongong, Australia. He earned an MSc in Mathematics from the University of Sherbrooke in Canada, an MSc in e-commerce from HEC Montreal, Canada, and a PhD in industrial engineering from the Polytechnic School of Montreal, Canada. His current research focuses on business value of IT, business analytics, big data, inter-organisational system (e.g. RFID technology) adoption and use, e-government (e.g. open data), supply chain management, electronic commerce and mobile commerce. He has published papers in a number of international conferences and journals including European Journal of Information Systems, Production Planning and Control, International Journal of Production Economics, Information Systems Frontiers, International Journal of Production Research, Business Process Management Journal, Proceedings of the IEEE, AMCIS, HICSS, ICIS, and PACIS. He is organizing special issues on IT related topics for the top IS and OM journals. Samuel Fosso Wamba is the corresponding author and can be contacted at: fossowam@gmail.com
Ngai Eric W.T. Eric W.T. Ngai eric.ngai@polyu.edu.hk Eric W.T. Ngai, PhD, is a Professor in the Department of Management and Marketing at The Hong Kong Polytechnic University. His current research interests are in the areas of e-commerce, supply chain management, decision support systems and RFID Technology and Applications. He has over 100 refereed international journal publications including MIS Quarterly, Journal of Operations Management, Decision Support Systems, IEEE Transactions on Systems, Man and Cybernetics, Production & Operations Management, and others. He is an Associate Editor of European Journal of Information Systems and Information & Management. He serves on editorial board of four international journals. Professor Ngai has attained an h-index of 20, and received 1,190 citations, ISI Web of Science.
Riggins Frederick Frederick Riggins fred.riggins@ndsu.edu Frederick Riggins, PhD, is an Associate Professor in the College of Business at the North Dakota State University. His research focuses on the Internet of Things, e-commerce, inter-organizational systems, RFID, and microfinance. He has published in leading journals including Management Science, Journal of Management Information Systems, Journal of the Association for Information Systems, International Journal of RF Technologies, Electronic Commerce Research and Applications, and Communications of the ACM. In a 2009 AIS publication, he ranked 9 on the list of top IS researchers from 2003 to 2007 based on number of publications and outlets. According to Google Scholar he has an h-index of 20 and over 3,300 citations. Akter Shahriar Shahriar Akter sakter@uow.edu.au Shahriar Akter, PhD, is a Senior Lecturer in the school of Management, Operations & Marketing at the University of Wollongong, Australia. Shahriar has published in leading international journals including Information & Management, International Journal of Production Economics, Electronic Markets, Journal of the American Society for Information Science & Technology, Behaviour & IT, International Journal of Production Research, Production Planning and Control, Communication of AIS, Journal of Selected Areas in Communication, etc. Shahriar's research interests include service systems evaluation, big data and business analytics, and complex modelling using PLS. Toulouse Business School, Toulouse, France Department of Management and Marketing, The Hong Kong Polytechnic University, Hong Kong, Hong Kong College of Business, North Dakota State University, Fargo, North Dakota, USA School of Management, Operations and Marketing, University of Wollongong, Wollongong, Australia . 2017. Guest editorial. *International Journal of Operations & Production Management* 37:1, 2-9. [Citation] [Full Text] [PDF]