Understanding Big Data Analytics capabilities in supply chain management: Unravelling the issues, challenges and implications for practice

Deepak Arunachalam [1], Niraj Kumar [2] and John Paul Kawalek [3]*

[1] *Sheffield University Management School, University of Sheffield, Conduit Road, Sheffield, S10 1FL, United Kingdom.*

[2] *University of Liverpool Management School, Chatham Street, Liverpool L69 7ZH, United Kingdom*

*Corresponding Author

Abstract

In the era of Big Data, many organisations have successfully leveraged Big Data Analytics (BDA) capabilities to improve their performance. However, past literature on BDA have put limited focus on understanding the capabilities required to extract value from big data. In this context, this paper aims to provide a systematic literature review of BDA capabilities in supply chain and develop the capabilities maturity model. The paper presents the bibliometric and thematic analysis of research papers from 2008 – 2016. This paper contributes in theorizing BDA capabilities in context of supply chain, and provides future direction of research in this field.

Keywords: Supply chain management, Big data analytics, Capabilities, Maturity model.

.

1. Introduction

The concurrence of events such as growth in adoption of supply chain technologies, data inundation and a shift in management focus from heuristics to data-driven decision making have collectively led to the rise of big data era. As a valuable asset for decision-making, Big Data Analytics (BDA) can play a pivotal role in transforming and improving the functions of supply chain. In this changing business environment, business leaders prefer to take decisions bearing in mind the data-driven insights rather than relying on their intuitions (Davenport, 2006). Due to the perceived benefits of BDA, organisations are highly motivated to develop their technical and organisational capabilities to extract value from data. Yet, practitioners face extreme difficulties in understanding the required capabilities to transform data into value. The core aspects of generating value depend on organisation's ability to capture, store and analyse a large volume of complex data generated in real or near real-time with the support of advanced analytics (Yesudas et al., 2014). Although the phenomenon of 'Big Data' is considered as the latest sensation worldwide, it has not essentially emerged impulsively. Last decade has seen a tremendous increase in adoption of a variety of Information and Communication Technologies (ICT) for Supply Chain Management (SCM), (e.g. RFID, Enterprise Resource Planning (ERP) to Internet of Things (IoT)). This has triggered huge data generation in the supply chain. Our continuous efforts to create more sophisticated technology to collect data at different stages of supply chain have resulted in the new era of big data.

The concept of supply chain signifies flow of information along with material and financial flow (Souza, 2014). Due to the adoption of ICT technologies, supply chains are enabled to monitor the information flow and inclined towards collecting and analysing a variety of data for efficient management (Chae and Olson, 2013). A typical supply chain has to manage the inflow of more than 100 gigabytes of data every day (The Economist, 2010). In fact, about 90% of data that are available today are generated by humankind in the last couple of years (Fawcett and Waller, 2014). It is estimated that the usage of RFID tags would increase rapidly to 209 billion units by 2021 (Marr, 2014; Tachizawa et al., 2015). The number of networked sensors used in automotive, retail and transportation have increased at the rate of 30 % per year (Manyika et al., 2011), with the perception that the sensor-based technology could substantially reduce the operational cost by 10 to 25%(Hahn and Packowski, 2015). The volume of digital data is growing exponentially and expected to reach 35 Zeta bytes by 2020 (Tien, 2015). In this current scenario, companies are increasingly

recognising the value of data and advanced analytics tools. BDA has the potential to govern Third Industrial Revolution (TIR), along with digital manufacturing, mass customisation and adaptive services (Tien, 2015). Adoption of BDA technologies could improve organisation capabilities in today's rapidly changing dynamic market environment (Meredith et al., 2012). However, to effectively deal with the diffusion of BDA technologies into the supply chain, organisational and behavioural issues related to adoption and practice has to be addressed. This paper finds very few empirical studies discussing the impact of BDA capabilities on SCM. Despite its popularity within the industry, many organisations are reluctant to invest in BDA technology due to the ambiguity in recognising potential benefits. Further, research on BDA capabilities in supply chain is very limited, and therefore, a comprehensive investigation of BDA capabilities is required to exploit the benefits of big data.

This paper adopts a systematic literature review approach to understand multiple dimensions of BDA capabilities in supply chain. Most of the academic studies have focused on exploring potential benefits of BDA from an investment perspective, but very little is known about key BDA capabilities for SCM. This paper begins with demystifying the concept of BDA in order to understand the evolution of big data. Then, through a structured search of literature between 2008-2016 with appropriate key words, 82 peer-reviewed journal papers and 13 maturity models are selected, to analyse and compile existing research in this domain; to define key BDA capabilities and maturity model specific to the supply chain context; and to address the challenges and issues of practising BDA. These research papers are critically reviewed to conceptualise the dimensions of BDA capabilities. Both bibliometric and thematic analysis techniques are used for the literature review. A maturity model is conceptualised to explain five dimensions of BDA capabilities, namely data generation capability, data integration and management capability, advanced analytics capability, data visualisation capability, and data-driven culture. Finally, the paper discusses the challenges and implications of BDA practice, and identifies the future direction of research in this field.

2. Business Intelligence (BI), Business Analytics (BA) and Big Data Analytics (BDA)

Data, according to Oxford English Dictionary (OED), is defined as "facts and statistics collected together for reference or analysis" or "the quantities, characters, or symbols on which operations are performed by a computer, which may be stored and transmitted in the form of electrical signals and recorded on magnetic, optical, or mechanical recording media." (Oxford Dictionary, 2016). In line with the first definition, data are

mainly facts and numbers collected for further analysis. The analysis of such data can be performed with less effort using traditional statistical methods and mathematical tools without the need of a computer. Whereas, the second definition encompasses 'characters' and 'symbols', emphasising the usage of the computer to perform analysis, and also the utilisation of additional medium to transfer and store data in various forms. Similarly, the term 'Big data' is defined by OED as "an extremely large data sets that may be analysed computationally to reveal patterns, trends, and associations, especially relating to human behaviour and interactions." These definitions have simplified the delineation and comparison of traditional data analysis with big data analysis. BDA uses more sophisticated computational techniques to handle complex data that has been increasing on a large scale, and unable to be processed using traditional methods.

Until recently, Business Intelligence (BI) was considered as an umbrella term which includes tools, techniques, and activities that converts raw data into useful information and support decision-making at different levels (Operational, tactical, and strategic) (Gudfinnsson et al., 2015). According to Sahay and Ranjan (2008), "BI refers to the use of technology to collect and effectively use information to improve business potency." Decision support, statistical analysis, data mining, forecasting, and OLAP are the key capabilities of BI, and four main components of BI are data sources, data marts, data warehouse and query and reporting tools (Sahay and Ranjan, 2008). Data warehouse is central to any BI solutions; data from internal and external sources are extracted and loaded into the data warehouse (Gudfinnsson et al., 2015). Another terminology that has been in practice is Business Analytics (BA), business users consider 'Business Analytics' as essential for providing data, information, and knowledge to support decision making (Acito and Khatri, 2014; Chen et al., 2012; Laursen and Thorlund, 2010), and its scope supposedly extends beyond traditional BI reporting. In literature, BI and Business Analytics (BA) terms are used interchangeably. Chae et al. (2014), referred BA as " the application of a broad range of analytical techniques and methods and data-driven analytic methodologies to different business domains." However, Laursen and Thorlund (2010, p.12) defined BA as "delivering the right decision support to the right people at the right time." Further, a related term called 'Decision Support Systems (DSS),' which arguably emerged during 1970's, also widely used in this context to indicate the usage of technological solutions to support decision-making problems (Shim et al., 2002). Bartlett (2013, as cited in Mortenson et al. (2015)) considers BI as an amalgamation of Business Analytics and Information Technology. Gudfinnsson et al. (2015) also supported

the argument of considering Business analytics (BA) as an integral part of BI. However, several researchers have a contrasting opinion and argued that BI is a division of analytics, and a new acronym 'BI&A' as a composite term was used (Chen et al., 2012). Mortenson et al. (2015) argued that there is a similarity between 'analytics' and 'Operations Research' or 'Management Science' as these concepts are related to improving operations and business decision-making. Supply Chain Analytics (SCA) is a context specific term commonly used to indicate Big Data and analytics activities in supply chain management (Wang et al. 2016; Sahay and Ranjan 2008; Souza 2014). Chae et al. (2014) argued that there are three sets of resources (Data management resources, IT-based supply chain planning resources, and Performance management resources) collectively constitute SCA. In general, supply chain analytics is the use of information and analytics tools to support efficient flow of material along the supply chain. Various definitions of SCA can be found in Rozados and Tjahjono (2014).

There is no clear consensus on different terminologies related to big data in literature. However, there is a pattern of evolution regarding the terminologies and development of capabilities suitable for data-driven decision-making. Chen et al. (2012) argued that the concept of first-generation BI (in 1990s) had evolved as a consequence of advanced statistical techniques in 1970's and data mining techniques in 1980's. While data mining was available for decades, it is commercially accepted only recently due to the data inundation problem and technological developments (Stefanovic, 2015). The second generation BI 2.0 has evolved during early 2000s (the period also witnessed the growth of the internet and web-based systems). Subsequently, BI 3.0 emerged in 2010's, leading us to the era of Big Data. The term 'Big Data' was primarily used by Cox and Ellsworth (1997) to refer to the storage challenges of datasets that are quite large and demands additional resources. The main distinction between traditional BI solutions and Big Data (BD) technologies is the scalability and ability to store a variety of data types (structured and unstructured) in real-time because the majority of traditional BI systems are inadequate and only able to store and analyse structured data aggregated at specific time intervals. Based on the literature, the evolution of BDA is illustrated in figure 1. Hence, BDA is not new, it has certainly evolved to meet the changing information processing needs of organisations. From 1950 to 2010, the complexity of data has increased gradually, and as a result, BDA has emerged as a flagship technology to tackle BD challenges.
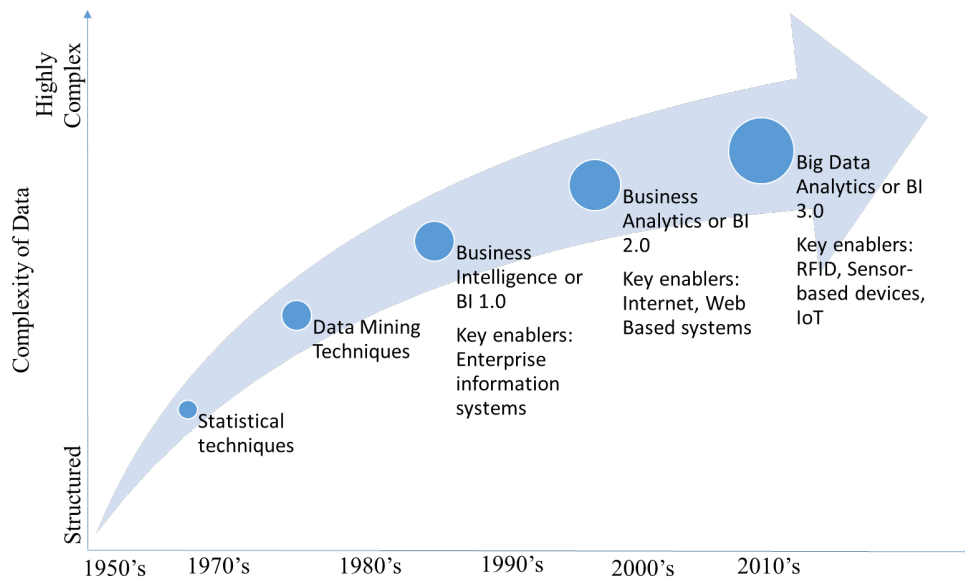
Figure 1: Evolution of Big Data Analytics

BD is often related to technological advancement and associated with new kinds of database architectures such as Hadoop, NoSQL, and distributed parallel processing of data. BD technologies are not only able to handle the volume of data but also can effectively manage a variety of data types such as textual information from online blogs, customer reviews, etc. (Mortenson et al., 2015). Vera-Baquero et al. (2015) argued that traditional BI systems are not process aware and insufficient to integrate data from heterogeneous data sources. A range of definitions of 'Big data' can be found in Wamba et al. (2015). A widely acceptable definition of BD from Gartner is given below:

"*high-volume, -velocity and -variety information assets that demand cost-effective, innovative forms of information processing for enhanced insight and decision making*" (Beyer and Laney, 2012, p.2).

In addition, according to Manyika et al. (2011), "Big data refers to datasets whose size is beyond the ability of typical database software tools to capture, store, manage, and analyse." Based on the nature of data, the BD was characterised mainly by three dimensions such as 'Volume,' 'Velocity,' and 'Variety' (Manyika et al., 2011; Sonka, 2014). However, apart from the 3V's, BD can also be characterised by another two dimensions 'Veracity' and 'Value' (Manyika et al., 2011; Neaga et al., 2015; Ge and Jackson, 2014). 'Volume' refers to magnitude of data generated; 'Variety' refers to "structural heterogeneity in a dataset" (Gandomi and Haider, 2015, p. 138); 'Velocity' refers to the speed at which data is

6

generated, analysed and acted upon (Gandomi and Haider, 2015); 'Veracity' or Verification refers to ensuring data quality, verifying unreliable and uncertain data; and 'Value' relates to economic benefits of Big Data (Mishra et al., 2016). Further, from the capabilities perspective, BDA is defined as "the capability to manage and analyse petabytes of data enable companies to deal with clusters of information that could have an impact on the business" (Hurwitz et al., 2013, p. 22). Wang et al. (2016 b) defined BDA from information lifecycle management view in the context of healthcare as "the ability to acquire, store, process and analyse a large amount of health data in various forms, and deliver meaningful information to users that allow them to discover business values and insights in a timely fashion (p.4)". Further, Hofmann (2015) defined the three V's of Big Data from capabilities perspective as "the ability of the organisations to process the volume of data integrated from a variety of data sources at a high velocity". However, recently Richey Jr et al. (2016) empirically explored the perception of SC managers about big data. While practitioners have recognised the Volume, Variety, Velocity and Veracity characteristics of Big Data, but there are no consensus attained regarding its definition. This paper defines BDA capabilities in context of SCM as,

"the ability of organisations to collect and organise supply chain data from heterogeneous systems distributed across organisational boundaries, analyse it either batch-wise or real-time or near real-time and visualise it intuitively to create proactive supply chain system and support decision making".

3. Research approach

While there are a few literature review papers (Mishra et al. (2016), Wamba et al. (2015), Donovan et al. (2015), Wamba and Akter (2015)) linking BDA and SCM, these reviews have not discussed BDA capabilities in supply chain context. Also, academic research and reviews related to BDA maturity models are scarce, and therefore this paper seeks to address these missing links. This review is aimed to summarise and describe existing research and conceptualise dimensions of BDA capabilities by synthesising the content of literature. This paper follows the literature review approach proposed by Mayring (2003). Similar approach is also used by Gao et al. (2016) and Seuring and Müller (2008) in analysing past research papers. This review approach includes four sequential steps: material collection, descriptive analysis, category selection and material evaluation.

3.1 Literature search

Section 2 demystified the concept of BDA and recognised several terms such as Big Data, Business analytics (BA), Business Intelligence (BI) and supply chain analytics. These terms are used as keywords to search literature in Big Data analytics in SCM. Unlike previous Systematic Literature Reviews (SLRs) in this domain, this study adopts a holistic approach by including all possible terms related to BDA practice in businesses. Similarly, different keywords related to SCM are also identified. Different combinations of terms are used to search relevant research papers. Scopus and Web of Science (WoS) databases are used to search related peer-reviewed papers for review. Table 1 summarises the keywords used for literature search along with the number of papers retrieved during initial search.

3.1.1 Inclusion and exclusion criteria

First, only journal papers published in English language are included. Consistent with Fahimnia et al. (2015), conference papers, papers in commercial magazines, and book chapters are excluded from the search to ensure quality, and only journal papers, reviews, and papers in the press are included. Although the initial search was not restricted with a time limit, the final shortlisted papers are published in the period of 2008- 2016. Initial shortlisting has produced 619 papers. After removing the duplication and verifying it in Endnote software, the full text of remaining papers are read to further eliminate irrelevant papers. Papers which clearly describe the application of BDA in SCM are only selected for the current review. This finally resulted in a total of 82 papers spanning from 2008 to 2016. The variety of search strings used in the SLR process is given in table 1, and the final list of 82 papers and 13 maturity models reviewed are given in Appendix A & B, respectively.

Next, the content of the selected papers is reviewed and classified based on categories such as the distribution of publication year, research methodology, among others. The analysis/evaluation process is complemented by the use of bibliometric analysis - to summarise existing research, and thematic analysis - to conceptualise the content of literature. For bibliometric analysis, BibExcel Software is used which requires meta-data information of selected journal papers in RIS format, and it is extracted from Scopus database. The findings from the descriptive analysis are given in section 4. Further, the contents of each shortlisted journal papers are analysed to find themes, codes and conceptualise dimensions of BDA capabilities maturity applicable to SCM context. For conceptualisation purpose, 13 maturity models are included in the review process.

Table 1: Initial search results

Search terms	SCOPUS	WoS
"Big Data" and "Supply chain"	104	65
"Big Data" and "logistics"	101	30
"Big Data" and "operations management"	15	8
"Big Data" and "operational performance"	1	1
"Big Data" and "operations research"	6	3
"Business Analytics" and "supply chain"	10	13
"Business Analytics " and "logistics"	6	4
"Business Analytics" and "operational performance"	2	2
"Business Analytics" and "operations management"	3	3
"Business Analytics" and "operations research"	4	5
"Business Intelligence" and "supply chain"	64	35
"Business Intelligence" and "logistics"	39	17
"Business Intelligence" and "operational performance"	7	4
"Business Intelligence" and "operations management"	3	3
"Business Intelligence" and "operations research"	6	3
"Business Intelligence" and "operational performance"	7	4
"Supply chain analytics"	8	4
"Supply chain" and "predictive analytics"	16	13

4. **Findings and discussion**

4.1 Distribution of papers per year

There is an increasing trend in terms of a number of papers published in the field of BDA in supply chain (Figure 2). In particular, a number of papers were published during 2015 and 2016 compared to previous years. Among 82 papers selected, most of them were published in last four years. The trend shows that there is growing importance among researchers to investigate the phenomenon of BDA in supply chain context.
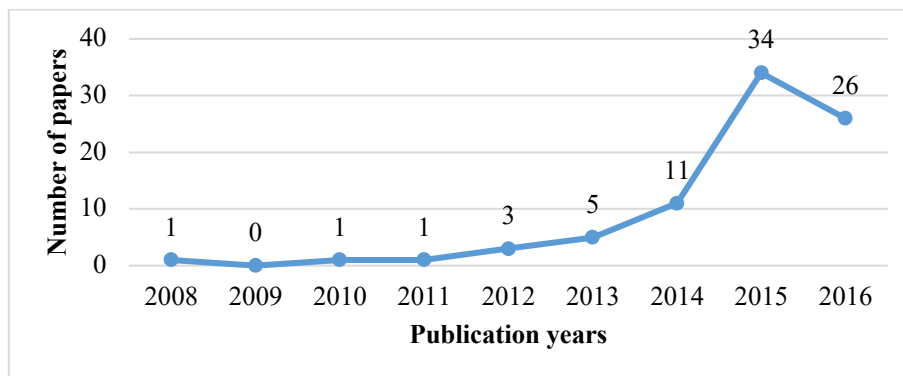

Figure 2: Distribution of research papers over the years

4.2 Top contributing authors

The top 10 contributing authors were extracted using BibExcel tool (figure 3). It is evident that Gunasekaran tops the list with 7 publications followed by Childe, Huang, Hazen,

Papadopoulos, Wamba and Zhong. However, the influence of authors was further evaluated using h-index and citation counts. It is found that Fawcett and Waller dominate in this criteria, followed by Chae and Gunasekaran (see figure 4).
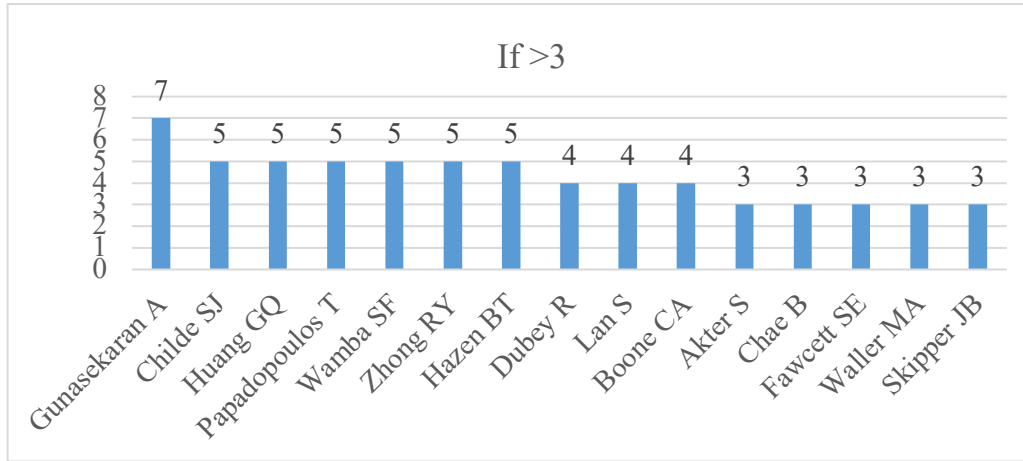

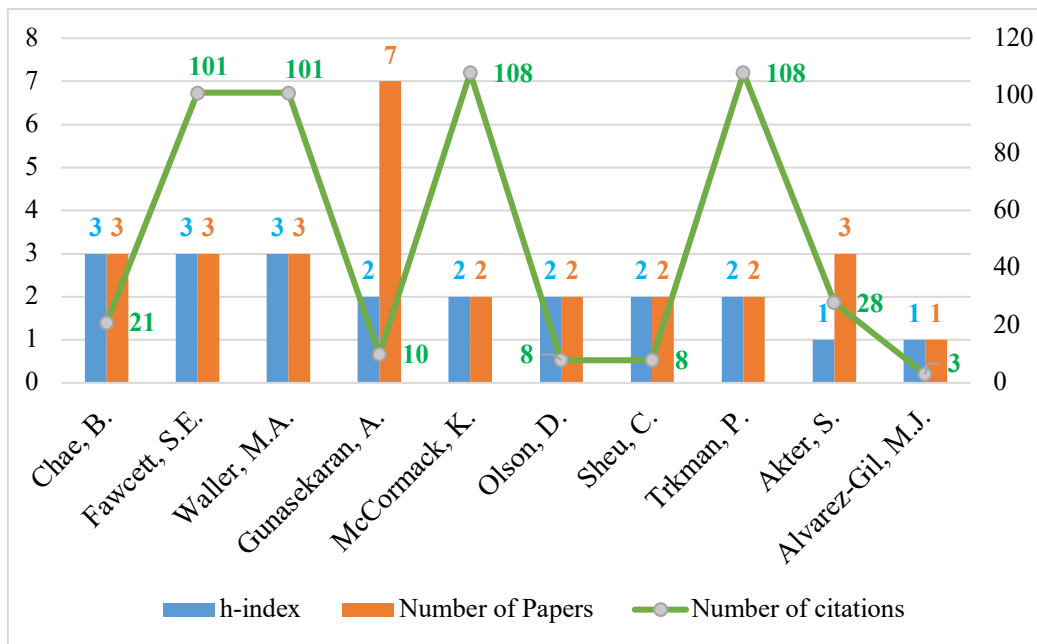*Figure 3: Top contributing authors based on number of publications*

Figure 4: Top contributing authors based on h-index, number of publications and citations

4.3 Affiliation statistics

The affiliations of all first authors are extracted using BibExcel. The United States dominated the top 5 list of contributing countries, which is followed by China, United Kingdom, India, and Germany.
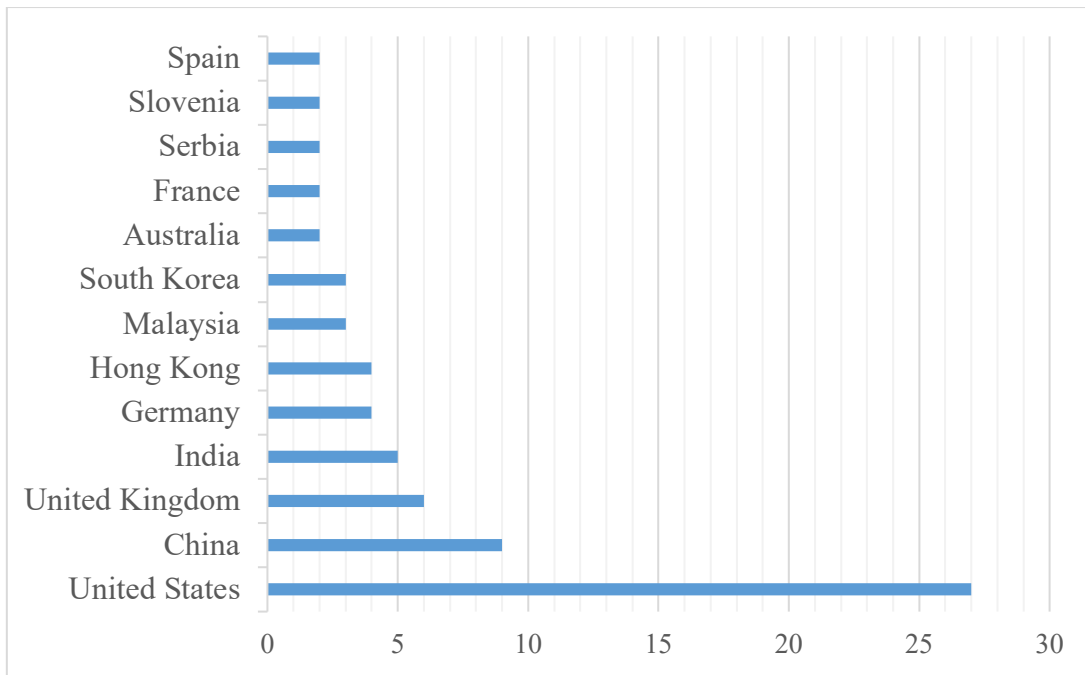
Figure 5: Classification of journal papers based on affiliation

4.4 Keyword statistics

Similarly, the keywords used in the selected journal papers are extracted using BibExcel, and the frequency of usage is analysed (Table 2). While the most frequently used keywords are 'Big Data' and 'Supply chain management,' it is evident that a wide range of other terms such as logistics, predictive analytics, RFID, and cloud computing are also used frequently.

Table 2: Keyword statistics

Keyword	Frequency	Keyword	Frequency
Big data	69	Logistics	10
Supply chain management	40	Internet of things	9
Supply chains	21	Radio Frequency Identification (RFID)	8
Manufacture	17	Big data analytics	7
Decision making	15	Supply chain	7
Data mining	14	Forecasting	7
Information management	12	Data handling	6
Data analytics	11	Cloud computing	6
Predictive analytics	11	Mass customization	4
Business analytics	10	Mass production	4

4.5 Theoretical Perspectives

The prominent theories used to explain the phenomenon of BDA adoption and practice are also reviewed. Figure 6 shows that Resource-Based View (RBV) is extensively used by 38% of research papers. Besides that, Dynamic capability theory (19%), Information processing view (15%), Contingency theory (8%), social capital theory (4%), and other theories are also applied frequently. Apart from that, frameworks such as Technology Organisation Environment (TOE) are also used to investigate organisations' BDA adoption behaviour(Chen et al., 2015). There is much scope for application and validation of several other theoretical lenses such as Knowledge based-view, absorptive capacity, systems theory, institutional isomorphism and agency theory to explore the phenomenon of BDA practice. Moreover, Hazen et al. (2016) have provided a review of several theories that can be applied in this domain.
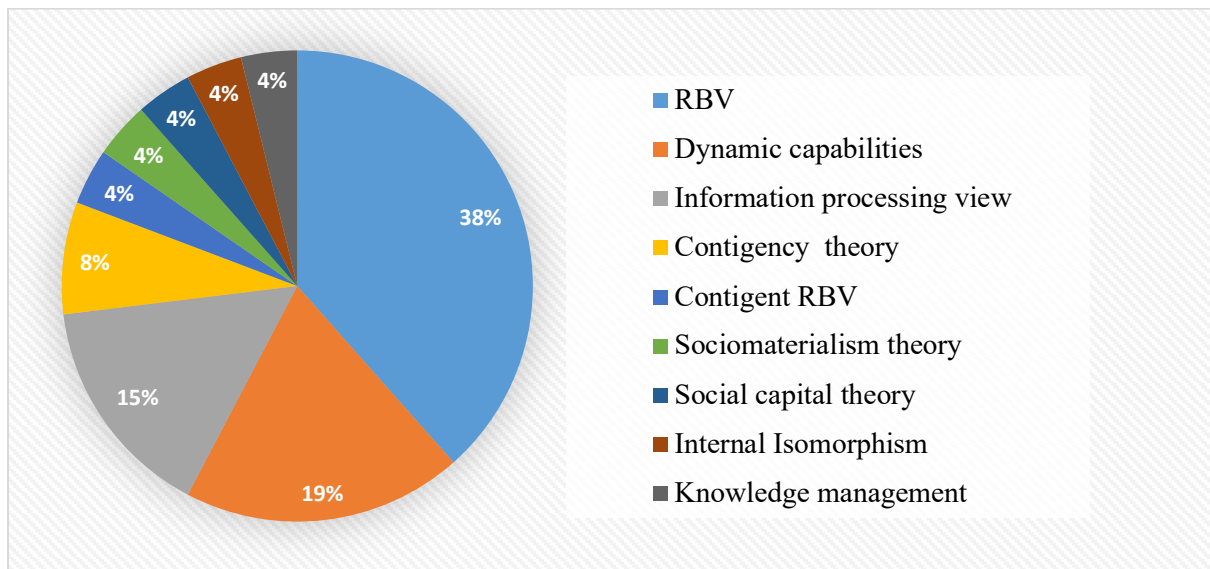

Figure 6: Frequency of underpinning theories in selected papers

4.6 Types of research methods

The research methods used by selected papers are also reviewed. Findings suggest that majority of studies have used conceptual approach, followed by case studies and Survey-based research (see Figure 7). Some papers have used experimental, analytical and mixed research approach. Interestingly, social media research has emerged as a new form of academic research method to address supply chain issues. In particular, recent studies by Bhattacharjya et al. (2016), Mishra and Singh (2016), Papadopoulos et al. (2016), Chae (2015) and Chan et al. (2015) have focused on the use of social media (i.e. Twitter and Facebook) data and text mining approach to address issues related to food supply chain, customer service and operations management. Likewise, case studies of integrating customer

reviews data from commercial websites and transactional data for better demand prediction is also observed in literature (Li et al. 2016, See-To and Ngai 2016). This change in research methods exhibits the influence of BDA on academic community as well.
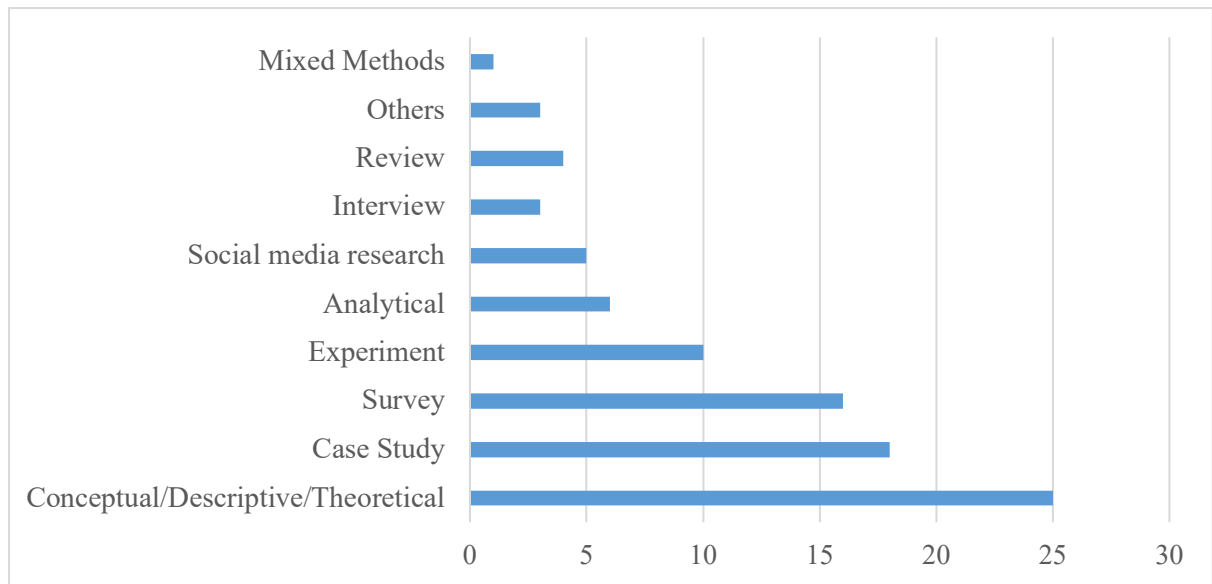

Figure 7: Classification of selected papers based on research methods

Moreover, among the survey-based studies, the distribution of papers based on the geographical representation of sample population is observed. Most of the studies have focused on global sample (n=4), followed by Iran (n=3) and other countries such as China, India, UK, Spain, and Germany. Several authors have recommended conducting studies in different geographical context and organisational culture (Bruque Cámara et al., 2015; Popovič et al., 2012). Moreover, 54% of survey studies have centered on manufacturing industry and remaining 46% of studies have focused on all types of sectors as research context. There is a lack of consideration and a particular emphasis on retail and other services sector regarding survey research.

4.7 Top cited papers

From the selected papers, Hu et.al. (2014), Trkman et al. (2010) and Waller and Fawcett (2013 b) have received the highest number of citations (Figure 8). Trkman et al. (2010) was the first to quantitatively investigate the impact of business analytics on supply chain performance, followed by several others such as Oliveira et al. (2012), Chae et al. (2014 a), Chae et al. (2014 b), Jamehshooran et al. (2015 a), Jamehshooran et at. (2015 b), Cao et al. (2015), Sangari and Razmi (2015), Chen et al. (2015) and Gunasekaran et al. (2017).
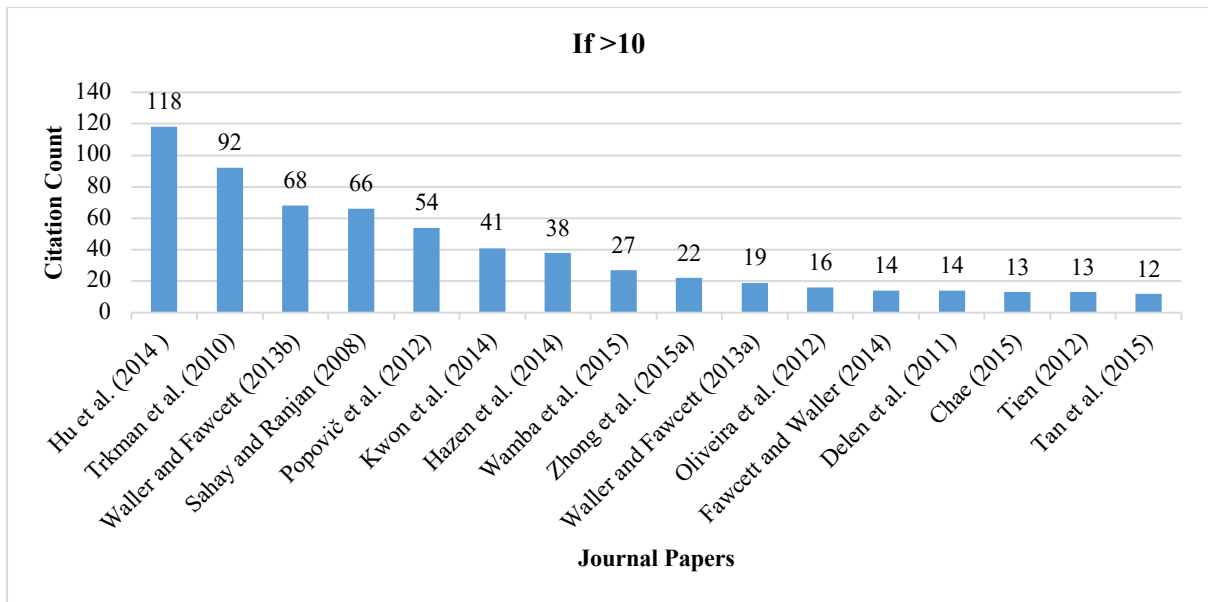
Figure 8: Distribution of papers based on number of citations

Lastly, Figure 9 illustrates the contribution of journals based on the number of research papers. Among the selected papers, International Journal of Production Research (IJPR), Annals of Operations Research (APR), International Journal of Production Economics (IJPE), Decision Support Systems (DSS), and Journal of Business Logistics (JBL) are the top 5 contributors in this domain. It is noticeable that most of these journals have contributed to the domain mainly during recent years, indicating acceptance of BDA as a thriving research in operations research / management community. Moreover, considering the growing interests it is expected that the figures would rise in coming years.
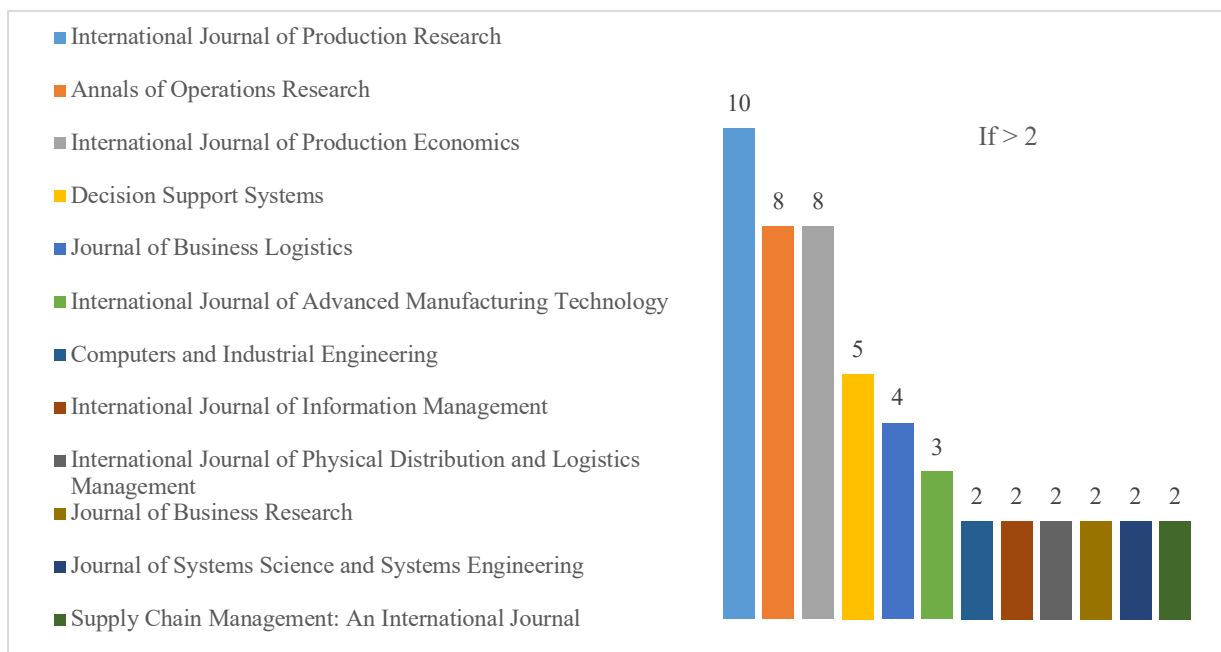

Figure 9: Contributions from different journals

5. Thematic Analysis

After the descriptive analysis, category selection and material evaluation are the third and fourth step in the review process towards conceptualisation of the concept. In this paper, the thematic analysis is used to explore BDA practice in SCM to conceptualise BDA capabilities. The conceptual framework and its dimensions have evolved during the analysis by identifying themes and coding it all along the review process. Since it is an emerging field, using a deductive approach would be inadequate because new codes have to be devised adaptably as it emerges from the data (Saunders et al., 2016, p. 582). Therefore, in this paper a bottom-up/data-driven approach progressing from individual dimensions to the whole by reading and rereading the selected papers, is used. The coding is verified independently by peer researchers to ensure reliability and validity (Seuring and Müller 2008). Minor amendments are made during the coding process to solve discrepancies between the authors. Finally, the papers are classified from capabilities perspective into five key dimensions; Data Generation (DG), Data Integration and Management (DIM), Advanced Analytics (AA), Data visualisation (DV) and, Data-Driven Culture (DDC). Moreover, literature also revealed the significance of cloud computing capability and Absorptive capacity as complementary to the key BDA capabilities.

5.1 BDA Capabilities in Supply Chain Context

This section presents the discussion around the capabilities identified from the thematic analysis. A conceptual framework is developed to simplify and delineate BDA capabilities identified (figure 10), and there are four quadrants in the framework; (i) Data poor and analytics poor, (ii) Data rich and analytics poor, (iii) Data poor and analytics rich, (iv) Data rich and analytics rich. The levels of BDA capabilities measured by 5 dimensions determine the quadrant within which organisations could find themselves. DG and DIM capabilities collectively constitute an organisation's level of data resources (data poor or data rich). Similarly, AA and DV capabilities denote an organisation's analytics resources (analytics poor or analytics rich). Data-driven culture is an intangible resource, fundamental to other capabilities, and has to be embedded into businesses routines related to BDA-driven decision making. Besides these, the cloud computing capability and Absorptive Capacity of organisations are considered as complementary resources and enablers of BDA. It can be argued that supply chain organisations need to possess all key capabilities to obtain value or wisdom from raw data as illustrated in figure 10. The stages of BDA assimilation are elaborated from assimilation and maturity perspective in section 5.2.
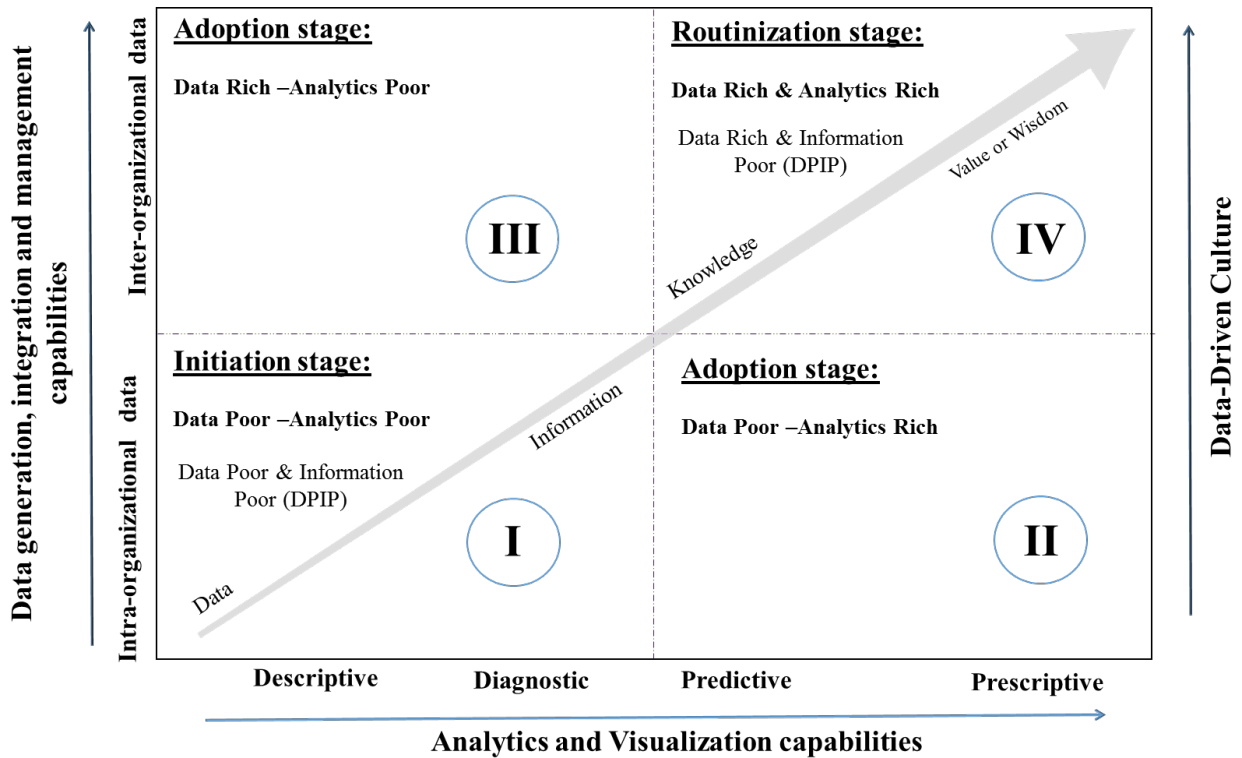
Figure 10: BDA capabilities framework for a supply chain

5.1.1 Data generation capability

Data Generation (DG) capability is the ability of organisations to seek, identify, create, and access data from heterogeneous data sources across organisational boundaries. If only there is a good quality data, insights can be drawn from it. DG capability facilitates the availability of BD to organisations disposal by establishing data sources, procedures and policies to generate required data for decision-making. In general, data can come from three different domains 'business, Internet, and scientific research' (Hu et al., 2014, p. 657). Adoption of supply chain technologies such as Advanced planning and scheduling (APS), RFID, ERP, CRM systems and Warehouse management systems (WMS) (Autry et al., 2010), are the primary sources and antecedents for the occurrence of data deluge in supply chains. The supply chain data generation was further revolutionised with the advent of the Internet of Things (IoT) technology facilitating real-time sensing and transfer of events data.

In context of a supply chain, data is scattered and can be acquired from diverse sources. The primary sources of data are from Enterprise information systems (EIS), which are mostly structured and transactional in nature. However, IoT, sensors, and RFID devices have the ability to convert the physical world into a virtual environment, which in turn generate a huge volume of unstructured data. Installation of RFID tags and readers on logistic

objects can convert them into 'passive smart logistics object' and 'active smart logistics object' (Zhong et al. 2015). When these physical objects start communicating via wireless communication, an enormous volume of data gets generated in real-time making it difficult to handle. For instance, a manufacturing plant that has deployed 1000 RFID readers and 10,000 tags could potentially generate Terabytes of data from a single day of operation(Zhong et al. 2015). It was reported that, by 2017, 50% of analytics implementation will utilise event-based streaming data (velocity) generated from machines or any physical objects that can communicate (Gartner, 2014). The data generated by event logs are also used by supply chain industries to perform process mining.

Adoption of electronic supply chain management (e-SCM) such as Internet-based inter-organizational systems, Internet-based electronic data interchange (EDI) has enhanced communication, coordination, and collaboration across organisational boundaries (Lin, 2014). The adoption of these systems would also generate volume of data through data exchange from the members of supply chain network. Supply chain practices such as collaborative planning, forecasting, and replenishment (CPFR) also generate additional data that needs to be stored and monitored (Chae and Olson, 2013). Apart from these, the ability to collect data from the internet and its significance in improving supply chain performance is also addressed recently (Bhattacharjya et al., 2016; Mishra and Singh, 2016).

Data is the raw material for analytics system and the data generation capabilities determines the position of an organisation on the continuum of data poor to data rich. Organisations that have leveraged data generation capabilities are supposedly in proximity to gain competitive advantage by means of meticulously sensing their business environment, collecting data and information and exploiting it for business growth. Organisations within a supply chain should constantly seek, identify and access data from heterogeneous data sources. Data generation capability indicates set of tools, technologies, and practices, which facilitate generation of huge volume, variety, and velocity of data. In order to generate right data for decision-making, organisations should put in place quality infrastructure and organisational practices to leverage the full potential of BDA. Based on the framework, it is argued that an organisation, which does not have the right infrastructure to support data generation and analytics, would fall into the quadrant of 'Data Poor and Information Poor'.

5.1.2 Data Integration and Management (DIM) capability

DIM capability is the ability of organisations to utilize tools and techniques to collect, integrate, transform and store data from heterogeneous data sources. The level of data integration, and ability to integrate different types of data gathered across organisational boundaries in real-time constitutes DIM capabilities.

Vast amount of supply chain data is distributed in heterogeneous sources and integration of these isolated data will be challenging (Stefanovic, 2014). DIM can be sub-divided into data acquisition from data sources, data pre-processing, and data storage. Due to the dynamic nature of supply chains, real-time access and scalability of data storage are the key capabilities to possess. Because of the complex nature of BD, the traditional database systems such as Relational Database Management Systems (RDBMS) are incompatible (Ge and Jackson, 2014). Chae and Olson (2013) considered that Inter-organisational systems (IOS) such as Web-based or cloud-based EDI (Electronic Data Interchange) could be used to enhance data integration capability. Using processes such as Extract Transform Load (ETL) or Extract Load Transform (ELT) data can be migrated into OLAP/OLTP systems. Data integration capability can improve visibility, responsiveness, and performance of material management, and provide a 360-degree view of manufacturing operations (Xiong et al., 2015). Wamba et al. (2015) have elaborated the importance of integrating Intra- and Inter-organizational data to improve service delivery via a representative case study. Tan et al. (2015) have exhibited the benefits of integrating data from multiple sources such as internal consumer data, social media data, and multimedia data to manufacture innovative products. Walmart integrates millions of transaction data generated every hour into one single system (Sanders, 2016). Moreover, to process data in real-time, in-memory databases and distributed or parallel processing approaches are appropriate than traditional approaches (Hahn and Packowski, 2015). Data storage component can include a set of hardware and software infrastructure that are suitable for scalability, storing data of different types that aggregate batch wise, real-time or in near real-time, and also systems to support fast querying and retrieval of data (Hu et al., 2014).

In addition, sharing of data generated from an entity to other parties in a supply chain could have a profound effect on optimising supply chain performance. Radke and Tseng (2015) elaborated the benefits of cloud computing for the purpose of data sharing in a more secure and agile manner within a supply chain network. Advanced big data technologies and cloud computing based data storage, distribution, and retrieval of information can enable

integration of data from various entities. Consequently, data acquired from heterogeneous systems has to be transformed into a standard format for further analysis. Data pre-processing activities such as ETL and ELT can be used to alter messy raw data into quality data sets suitable for analytics operations. However, data integration capability has to deal with two main issues data inaccuracy and redundancy (Chae and Olson, 2013). BI vendors suggest organisations put emphasis on data governance too as it is the biggest challenge that an organisation could face (Meredith et al., 2012).

In supply chain context, there could be several reasons for reluctance in sharing data among its members. Some of the major barriers to data sharing are sensitivity, fear of losing competitive advantage and problems with access and control of information sharing (Radke and Tseng, 2015). Supply chains are dynamic in nature and prone to environmental uncertainty. So, to streamline data sharing process, it is imperative to have data access control with firms in a short relationship and limit the access when the relationship perishes. If data is scattered and not integrated to provide a single point of truth, organisations would lack a valuable resource for decision making i.e. quality data; therefore, SCM demands DIM capabilities.

5.1.3 Advanced analytics capability

Analysing Big Data is a complex task and requires advanced analytics capabilities. *Advanced Analytics capabilities* is defined as the ability of organisations to utilise tools and techniques to analyse supply chain data in batch wise, real-time, near-time, or as it flows and extracts meaningful insights for decision making. Data analytics is the most significant phase in data value chain from raw data to meaningful insights; analytical tools and techniques are leveraged to slice through the data to data-driven insights. Depending on the depth of analysis, data analytics techniques are classified into descriptive, predictive and prescriptive analytics (Souza, 2014). As shown in figure 10, with the increase in analytics capabilities from descriptive to prescriptive analytics, organisations decision-making capabilities at operational (short-term), tactical (mid-term) and strategic (long-term) level would certainly improve. Descriptive analytics is based truly on the principle of classical statistics methods. But, predictive analytics is the combination of statistics, data mining, and machine learning techniques (Blackburn et al., 2015). While descriptive analytics rely on historical data, predictive analytics can utilise historical data and dynamically can include external data from other sources but contingent upon analytics models and algorithms used. Manyika et al. (2011) discussed several analytics techniques available to analyse Big Data such as

association rule mining, classification, clustering, crowdsourcing (data collected through web 2.0), data fusion and integration, ensemble learning, machine learning, genetic algorithm, network analysis, neural network, predictive modelling, sentiment analysis, spatial analysis, and time series analysis.

Due to the dynamic nature, SCM demands real-time data analytics capabilities, which enables processing of real-time data (such as RFID data) to monitor processes and events (Manyika et al., 2011). Hahn and Packowski (2015) reviewed in-memory analytics application for real-time data processing in supply chain context, and categorised its application into (i) Plan and optimise (ii) Predict and act, (iii) Sense and respond, (iv) Monitor and navigate. Accordingly, descriptive analytics on RFID and GPS data would enable continuous monitoring of material flow in supply chain, increasing the visibility of assets. Predictive analytics aims to identify patterns in data and converts it into business rules. It can help organisations to sense and react to supply risk, demand risk, and operations risk, and also can be used to predict delivery, asset utilisation, maintenance, productivity, and sales forecast. For instance, data mining techniques are applied in the area of supply chain fraud detection to reducing supply chain risks (Kraus and Valverde, 2014). Similarly, leveraging prescriptive analytics can be beneficial to plan and optimise supply chains either at operational or strategic level using aggregate data from transactions and granular data from day to day business operations.

Chen et al. (2015) have conceptualised the use of BDA in SCM into three categories; (i) Coordination/Integration process (Warehouse operations improvements, Process/equipment monitoring, and logistics improvements), (ii) Learning processes (sourcing analysis, purchasing spend analytics, CRM/customer/patient analysis, forecasting/demand management - S&OP, and Inventory optimization), (iii) Reconfiguration processes (network design/optimization, production run optimization, inventory optimization). An extensive multistage study conducted by Sanders (2016) illustrates real-world BDA applications in various areas of SCM including but not limited to inventory optimisation, labour scheduling, route optimisation, price optimization and micro-segmentation in marketing. Moreover, the accuracy of demand forecasting, one of the critical aspect of SCM, can be improved using advanced predictive analytics techniques outperforming historical data based statistical techniques (Blackburn et al., 2015). Predictive analytics capability enables organisations to consider both endogenous and exogenous variables while forecasting demand. The unstructured customer reviews have variables that

can predict sales nowcasting (See-To and Ngai, 2016). Traditional forecasting depends on aggregated data, but by deploying real-time analytics capabilities organisations can analyse demand data in real-time increasing accuracy and can reduce bullwhip effect (Hofmann, 2015). Increasing robustness of demand forecasting via predictive real-time analytics can eventually improve other functions such as production planning and inventory optimisation which rely on forecasted demand. Moreover, since a huge volume of spatiotemporal data is generated from GPS and RFID devices, predictive and spatiotemporal analytics can be used to analyse these unique data types, for instance, to predict truck arrival time (van der Spoel et al., 2015) and optimising blood supply chain (Delen et al., 2011). Information access and content quality is empirically found to increase by leveraging data integration and analytics capabilities (Popovič et al., 2012). It is also found to increase organisations information processing capabilities (Cao et al., 2015) and supply chain planning satisfaction (Chae et al., 2014a).

Tan et al. (2015) described how text analytics capability could be combined with other analytics capabilities to develop innovative products. Markham et al. (2015) also emphasised the significance of text analytics capabilities to make new product development decisions. It is also suggested that integration of twitter data mining capabilities into CRM systems would be beneficial to resolve queries in real-time and improve customer satisfaction (Bhattacharjya et al., 2016). Moreover, applying predictive analytics (decision tree algorithm) on unstructured data from commercial websites, which include transactional as well as behavioural data of consumers, can help e-commerce supply chain to identify predict and manage demand chain (Li et al., 2016). Despite its potential impact on mainstream SCM areas of planning, sourcing, making and delivering of products, Boone et al. (2016) claimed that BDA could be beneficial to enhance the performance of after-sales service parts management. Examples of BDA application and its benefits discussed here are just tip of an iceberg. It has enormous potential to transform traditional reactive supply chains into proactive data-driven supply chains. Hence, as argued by Markham et al. (2015), data collection (Big data) is only the first step in data value chain, and advanced analytics capabilities are required to extract value from it.

5.1.4 Data Visualisation capability

Data Visualisation capability is the ability of organisations to utilise tools and techniques to render information visuals and deliver the data-driven insights intuitively in a timely manner to the decision makers. Data visualisation is "the representation and

presentation of data that exploits our visual perception abilities in order to amplify cognition" (Kirk, 2012, p. 17). The two main purpose of data visualisation, which is a form of descriptive analytics, is sense-making (data analysis) and communication of abstract information through the graphical display (Few, 2009).

In supply chain context, data visualisation capabilities are equally important compared to other BDA capabilities. It can be beneficial for SCM in various ways such as monitoring future demand, optimise distribution networks, visualise delivery routes and material tracking. The primary goal of data visualisation is to intuitively represent knowledge and information using various techniques such as tag cloud, graphs, Clustergram, and heat maps (Chen and Zhang, 2014; Manyika et al., 2011). In particular, heat maps which could display geographical information such as consumption location, transaction density, etc. is valuable to develop new distribution strategies (Tachizawa et al., 2015). Tools such as desktop applications, interactive web and mobile applications offer capabilities to executives, decision makers and customers to interact with the analytics ecosystem (Barlow, 2013). The presence of data visualisation capabilities can effectively support the three analytics capabilities (descriptive, predictive, and prescriptive) (Assunção et al., 2015). From the usability perspective, highly interactive visualisation would help users to interpret, find patterns and make decisions quickly and effectively from raw data as well as content from analytics system.

Manyika et al. (2011) emphasised that insight should be presented in such a way that people can effectively consume it and aid them to take action. Data visualisation tools are significant to monitor key performance indicators (Xiong et al., 2015). eBay leveraged Tableau (a popular data visualisation platform) to transform large complex data sets into interesting pictures for decision makers to understand consumer behaviour and trends via techniques such as sentiment analysis (Chen and Zhang, 2014). Park et al. (2016) have developed a visual analytics based decision support system (DSS) incorporating predictive analytics capabilities and experimented with supply chain network data. They argued that interactive visualisation would enhance human cognition level while decision-making. Similarly, GIS-based analytics and visualisation capabilities are found to be beneficial to effectively manage blood supply chain (Delen et al., 2011). Zhang et al. (2013) used data visualisation techniques to identify sources of contamination in food supply chain. Brandau and Tolujevs (2013) experimented with visualisation techniques and clustering algorithms to manage irregularities in real-time sensor data and improve logistics performance. Moreover,

Bock and Isik (2015) have justified the importance of visualisation in informing managers about the current status of processes.

The paper argues that the ability to create and transfer the visual representation of analytics insights in real-time reduces the time difference between the process of data collection, storage, analytics, and reporting. For SCM, integrating the large volume of data and visualising actionable insights should be the priority to develop operational intelligence solutions (Yesudas et al., 2014). In fact, it is widely accepted that many organisations prefer to use intuitive data visualisation tools (such as Tableau, Spotfire, QlikView, Datameer, etc.) that are catered to the needs of information visualisation problems (Zhong et al., 2016 b).

5.1.5 Data-Driven Culture

Data-driven culture is an intangible resource that represents the beliefs, attitudes, and opinion of people towards data-driven decision-making. According to Aho (2015, p.284) "The transformative potential of Big Data lies in treating data as an asset." Aho (2015) argued that Big Data involve extensive change management and development of new organisational culture to transform the organisation. According to Kiron and Shockley (2011), three main characteristics constitutes data-driven culture they are, (i) Analytics should be treated as a valuable or strategic assets by an organisation, (ii) Top Management should support and leverage analytics across the organisation, and (iii) Accessibility of Data-Driven insights to decision makers. Real-world case examples have suggested engagement of implementation team and top management support are significant for developing BDA capabilities (Wamba et al., 2015). Lavalle et al. (2010) conducted a survey to identify the significant barriers to adoption of analytics in organisations. Despite the challenges of technology and data quality, organisations considered culture and managerial issues as the significant barrier to analytics adoption. Sangari and Razmi (2015) have empirically verified that organisation find it difficult to establish cultural competence in supply chains. To build transformational BDA capabilities, business and IT leaders in organisations have to work together and develop new strategies and roles such as Chief Data Officer, Data scientists to address the needs of technology and business. Moreover, Cao et al. (2015) argued that presence of the data-driven culture would facilitate organisations to make data-driven decision and rely on fact-based decision to develop new products and services. Certainly, organisations who possess advanced analytics capabilities could not extract full value if not effectively integrated into the business decision-making process and not accepted as a decision-making tool (Blackburn et al., 2015). Not only within an organisation, but it is also

imperative to develop data-driven culture across a supply chain. For example, Walmart extends their analytics capabilities to all of its suppliers and promotes the culture of data-driven decision-making across their supply chain to improve performance.

5.1.6 Cloud computing capability

Integrating cloud computing capability to leverage BDA for SCM can be a way forward to attain maximum value from the investment, as it has potential reduce some of the challenges discussed in section 6. Cloud computing infrastructure can support the integration of data from both intra- and inter-organisational sources and provide scalable storage. From Resource Based View perspective, cloud computing can be seen as a complementary resource while BDA being the primary resource for improving performance.

Cloud computing is considered as a key technology enabler of BDA capabilities. Neaga et al. (2015) have argued that cloud-enabled big data systems could improve logistics performance, information and knowledge sharing with all stakeholders in logistic service and increase customer satisfaction via proactive service. Different types of cloud-based service models are available such as infrastructure as a service (IaaS), platform as a service (PaaS) and Software as a service (SaaS), and these services can be further deployed in four different ways, private cloud, community cloud, public cloud, and hybrid cloud (Bruque Cámara et al., 2015). In the context of a supply chain, cloud-computing services such as IaaS can be deployed on a community level to facilitate data sharing between supply chain partners within the network. Neaga et al. (2015) argued that several business models based on cloud computing such as Analytics as a service (AaaS), Big Data as a service (BDaaS) and Knowledge and information as a service (KaaS/IaaS) can be adopted to leverage BDA. Further, cloud computing can be used to integrate and store structured and unstructured data from different sources, along with cloud-based data warehousing and Cloud-based data sharing (Tien, 2012). Alfaro et al. (2015) argued that to facilitate access of the analysed data and visualise it into cloud, an integrated architecture of personal cloud, supply cloud, manufacturing cloud, warehouse cloud and logistics cloud can be utilised to transfer information in near real time. Integrating BDA on cloud will improve virtual integration causing supply chain boundaries to disappear(Fawcett and Waller, 2014). Singh et al. (2015) developed a cloud-based framework that can acquire data from all members (from farmers to retailers) in a beef supply chain. The application of cloud-enabled BDA can be used to monitor carbon footprint of supply chains (Singh et al., 2015), and visualise complex RFID data (Zhong et al., 2016 a). Similarly, 'SCEnAT,' an environmental life cycle assessment

tool, leverages cloud-enabled BDA capabilities to analyse data and intuitively visualise carbon hotspots across the supply chain (Koh et al., 2013).

5.1.7 Absorptive capacity

Absorptive Capacity (ACAP) as an organisational capability can play a significant role in both assimilation and extraction of value from BDA. Cohen and Levinthal (1990, p.128) defined Absorptive capacity (ACAP) as "the ability of a firm to recognize the value of new, external information, assimilate it, and apply it to commercial ends is critical to its innovative capabilities." Malhotra et al. (2005, p.145) perceive it as "the set of organizational routines and processes by which organizations acquire, assimilate, transform, and exploit knowledge to produce dynamic organizational capabilities." Moreover, Zahra and George (2002), reconceptualised ACAP into potential ACAP (PACAP) and Realised ACAP (RACAP). According to Roberts et al. (2012), in the context of technology assimilation, ACAP is treated as an asset in the form of prior knowledge possessed by organisations which foster innovation. BDA as a knowledge infrastructure could enhance knowledge transfer from supply chain partners and increase recipient firm's ACAP. On the other hand, in relevance to extracting value from technology like BDA, ACAP can be conceptualised as a dynamic capability, which could complement BDA capability in generating business value.

ACAP is used by many researchers to explain organisational learning from a strategic management perspective. ACAP is a multi-level and multi-dimensional construct. It relates to individual level to inter-organisational level (Roberts et al., 2012), and can have many interrelated capabilities (Wang et al., 2014). In supply chain context, the critical information needed to improve supply chain performance is mostly available in external sources (Dobrzykowski et al., 2015), not readily accessible for decision-making. However, BDA can provide that critical information in real-time and enhance the organisational capability to acquire, assimilate, transform, and exploit the information and knowledge for commercial ends. On the contrary, firms with low absorptive capacity would find it difficult to adopt innovative BDA technologies such as MapReduce (Ebner et al., 2014). Moreover, it can be argued that even if BDA resources are well established at the organisation level, it becomes obsolete when an organisation does not exhibit absorptive capacity. In fact, ACAP is considered as one of the prerequisites of BD initiative for successful implementation (Kabir and Carayannis, 2013).

5.2 BDA Maturity in Supply Chain Context

Organisations who are better informed of their current state of BDA capabilities are in better position to tackle challenges while adoption (Lavalle et al., 2010). Maturity models (MM) can help organisations to measure their current state of technological competence benchmarked against industry standards. The term 'maturity' was defined as "state of being complete, perfect, or ready" (Simpson and Weiner, 1989). Starting from Nolan (1973), several BI and Big Data MM were developed both in academia and industry to help organisations realise their current potential and prepare for future endeavours. Narrowing the literature review to MM specific to BDA, a few related studies (Brooks et al., 2015; Cosic et al., 2015; Gudfinnsson et al., 2015; Spruit and Sacu, 2015) are identified. However, models from non-academic sources are also considered and finally 13 maturity models are reviewed in this paper (*see Appendix B*). The models are evaluated based on several criteria such as dimensions used, number of maturity stages, transparency of methods used to develop and test the models, acceptance level by business and academic community, and the availability of assessment tools. For instance, the maturity model developed by Lavalle et al. (2010) contains three levels of maturity (aspirational, experienced, and transformed) which was used by Gudfinnsson et al. (2015) to evaluate a company's BI maturity using a case study approach. Popovič et al. (2012) conceptualised that BI systems maturity composed of two dimensions: Data integration, and Analytical capabilities. Most of the MM are generic and there is a need for context-specific BI or BDA MM (Brooks et al., 2015). Thus, integrating the dimensions of BDA capabilities and MM (see Table 3), a BDA capabilities maturity model specific to the context of SCM is conceptualised in this paper. The two significant aspects of maturity models are its measurement dimensions and stages of maturity. The stages of maturity are discussed here using the proposed framework (Figure 10).

In this paper, drawing upon the theory of diffusion of innovation and stages of assimilation (Fichman and Kemerer, 1997), the assimilation of BDA is conceptualised into four stages, incognizant stage, initiation stage, adoption stage and routinisation stage. Assimilation refers to the extent of technology defuses and routinised into business processes (Roberts et al. 2012). At the incognizant stage, firms would lack knowledge about BDA and its benefits. At initiation stage, firms are aware of BDA and considering it for leveraging, but have not implemented it. Organisations that are categorised into incognizant and initiation stages (first quadrant) are presumably Data Poor and Information Poor (DPIP), as the level of BDA capabilities will be low. Whereas organisations in adoption stage (second and third

quadrant) are aware of BDA technology, and involve in the process of adopting it. However, depending on the nature of capabilities they currently possess, an organisation in adoption stage can further be categorised into; (i) Organisation who possess a high level of Data Generation and Data Integration and Management Capabilities but not Advanced Analytics capabilities as Data Rich and Information Poor (DRIP), and (ii) Organisations that have leveraged some form of analytics capabilities, but not collecting and integrating data from internal and external sources to the fullest. Besides, organisations that possess a high level of all key BDA capabilities and fully integrate their business processes are at the stage of routinisation. These organisations will be the leaders in BDA practice and are certainly Data Rich and Information Rich (DRIR). Overall, assimilation of BDA capabilities can positively influence supply chain performance but it requires commitment from top management (Gunasekaran et al., 2017).

Table 3: Dimensions of BDA capabilities maturity

Dimensions of BDA capabilities maturity	Key elements	Key references
Data generation capability	1) Data generation infrastructures. 2) Data sources. 3) Repository for open data. 4) Strategies to collect data from sensors and other devices in real-time. 5) Data gathering capability.	Brandau and Tolujevs (2013); Zhang et al. (2013); Jin and Ji (2013); Hu et al. (2014); Radcliffe (2014); Janssen et al.(2014).
Data integration and management capabilities	1) Data integration from heterogeneous sources. 2) Data warehouse capability 3) Data management resources (capture, manipulation and redistribution of data) 4) Data governance 5) Unstructured data management 6) Data quality (accessibility, completeness, timeliness, reliability, consistency, and accuracy) 7) Metadata repository	Lavalle et al. (2010); Popovič et al.(2012); IDC (2013); Chae et al. (2014a); Chae et al.(2014b); Halper and Krishnan (2014); Nott (2014); Knowledgent (2014); Radcliffe (2014); Spruit and Sacu (2015); Sulaiman et al. (2015); Cosic et al. (2015); Tan et al.(2015).
Advanced analytics capabilities	1) Analytics in the plan, source, make and deliver. 2) Supply chain analytics (1) Sustainable SCA, 2) Agile SCA, 3) Collaborative SCA, 4) Process-based SCA, 5) Functional SCA). 3) Predictive analytics. 4) Real-time analytics. 5) In-memory analytics 6) Data mining (time series analysis, association rule mining, classification and clustering analysis). 7) Web mining and text mining. 8) Online Analytical Processing (OLAP). 9) Trend analysis and "What-if" scenario analysis. 10) Data analysis and Data decision capability.	Lavalle et al.(2010); Popovič et al.(2012); Trkman et al. (2012); Oliveira et al. (2012); Zhang et al. (2013); Brandau and Tolujevs (2013); Hu et al.(2014); Nott (2014); Radcliffe (2014); Knowledgent(2014); Sulaiman et al.(2015); Blackburn et al.(2015); van der Spoel et al.(2015); Wang et al.(2016a).

| Data visualisation capabilities | 1) Interactive visualisation.
2) Dashboards and key performance indicators (KPI).
3) Real-time information monitoring.
4) Strategic and operational reporting using historical and streaming data. | Popovič et al. (2012); Tien(2012); Brandau and Tolujevs (2013); Radcliffe (2014); Yesudas et al. (2014); Dutta and Bose (2015); Neaga et al. (2015); Sulaiman et al.(2015); See-To and Ngai(2016). |
| Data-driven culture | 1) Cultural and political issues.
2) Culture and execution.
3) People (Skill, Culture and Organisation structure).
4) Culture Capability.
5) Cultural competence in the context of supply chain.
6) Data and analytics as a valuable asset for decision-making. | IDC(2013); Halper and Krishnan(2014); Nott(2014); Cosic et al.(2015); Dutta and Bose (2015); Sangari and Razmi(2015). |

6. Issues and challenges of adopting and practising BDA

6.1 Organisational challenges

a) *Time-consuming:* Predictive analytics initiative is time-consuming and includes various stages of developing, testing and adapting it to different contexts (Blackburn et al. 2015). Bringing together experts from various functions with varied mindsets will be a challenging task. In complex systems like supply chain, BDA implementation needs consistent support from top management and key stakeholders, as it might take 12-18 months to see results. Getting access to data (owned by different departments of the organisation), combining, validating, and data cleansing will be a tedious which requires exhaustive commitment from the project management team.

b) *Insufficient resources:* The data and analytics resource capabilities vary across firms in a supply chain network. Supply chain partners' lack of IT resources and capability to share data and information in real-time will cause discrepancies. As stated in Dutta and Bose (2015), collaboration and cross-functional team formation between various stakeholders within an organisation should be a priority for implementation of big data. However, while forming an inter-organisational cross-functional team, the challenges that can be anticipated are competition within supply chain network, principle-agent conflicts, incentives arrangements, data sharing policies, etc. Data-driven culture, which is one of the key BDA capabilities, and fact-based management, should be encouraged across supply chain network as a strategy for effective utilisation of BDA and to create business value. Leadership also plays a significant role in successful implementation of BDA systems for SCM (Seah et al., 2010).

c) *Privacy and security concerns:* Big Data possess several concerns such as privacy, security, unethical use of Big Data and processing data ineffectively (Hu et al., 2014),

which would lead to biased findings (Tien, 2012). Supply chain professionals raised concern about privacy and data security, and argued that out-dated regulations are one of the major obstacles in data sharing, especially consumer data (Richey Jr et al., 2016). Privacy, security and data laws could be of serious concern for multinational supply chains, obligated to abide by the laws of different countries while sharing data across supply chains (Alfaro et al., 2015). However, these challenges could be overcome by employing effective data governance initiative within the process of data integration and management.

d) *Behavioural issues:* From the behavioural perspective, the use of real-time data and information could be challenging because decision makers may excessively react to even small changes in the physical world which would worsen the "bullwhip effect" and increase supply chain risk and cost of inventory (Tachizawa et al., 2015). Tachizawa et al (2015) argued that concerning Big Data there is a risk of identifying many statistically significant but irrelevant correlations that do not have a causal linkage. Big Data focuses predominantly on correlation but not on causation, which necessitates human inferences to solve big problems. Nevertheless, Google has proposed a heuristic approach to solve the correlation and causality problems (Radke and Tseng 2015).

e) *Issues with Return on Investment (ROI)*: Unclear benefits and ambiguity on ROI make stakeholders apprehensive about implementing BDA (Richey Jr et al., 2016; Sanders, 2016). Achieving financial benefits from BDA is challenging too as it depends mostly on the "downstream" employees who performs the task (Davenport et al., 2001). For instance, analytics can help segment the market based on available data, but it is the sales/marketing team who has to believe in the data-driven insights and treat customers based on the segment types to make real change. Here, Data-driven culture and the employees' absorptive capacity at individual level plays' a crucial role in absorbing and assimilating the knowledge.

f) *Lack of skills:* Schoenherr and Speier-Pero (2015) identified potential barriers to using predictive analytics in SCM through a survey. The primary barriers are inexperienced employees, time constraints, lack of integration, lack of appropriate predictive analytics solution, and issues with change management. Moreover, it was found out that professionals who plan to use analytics in future and currently not using it have considered the lack of data and inability to identify suitable data as a prominent barrier, which relates to the situation of 'Data Poor and Information Poor'.

Waller and Fawcett (2013 b) argued that data scientist requires a combination of both analytical skills and domain knowledge, which is difficult to find such combination as someone good in analytical skills may not be interested in learning domain knowledge. A recent study has also confirmed that lack of experts in BDA is a serious issue among supply chain professional (Richey Jr et al., 2016).

6.2 Technical challenges

a) *Data scalability:* Richey Jr et al. (2016) identify data scalability as a major technical issue in BDA adoption. Organisations have to dump their data after a particular period so as to store newly generated data. Replacing relational databases which are limited regarding scalability with more advanced infrastructure such as Hadoop distributed databases, distributed file systems, parallel computing and cloud computing capability could be considered to tackle scalability issues. NoSQL database which has a high level of scalability is a better choice to deal with unstructured data generated from IoT data sources (Kang et al., 2016). However, leveraging cloud-computing capability to store BD could incur more financial burden to organisations as with increase generation of BD cloud storage utilisation cost will also eventually increase. To avoid this, organisations could adopt strategies to optimise data collection process and reduce unwanted data generation right from the source (Rehman et al., 2016)*.*

b) *Data Quality*: Supply chain managers rely on data-driven insights for various reasons such as to gain visibility, collaboration, process control, monitoring, optimisation, etc. and, ultimately aim to obtain competitive advantage(Davenport, 2006; Hazen et al., 2014). However, there are quality issues associated with the process of data production, which is often compared with the product manufacturing process (Hazen et al., 2014; Wang, 1998; Wang et al., 1995; Wang and Strong, 1996). Hazen et al. (2014) stated that poor data quality would hinder the data analytics activities and affect management decisions. Unlike a physical product, data is intangible in nature and measuring data quality is a multidimensional problem (Hazen et al., 2014). Hazen et al. (2014) discussed the works of Wang and Strong (1996) and Lee et al. (2002), who classified the dimensions of data quality into *intrinsic* (accuracy, timeliness, consistency, and completeness) and *contextual* (relevancy, value-added, quantity, believe-ability, accessibility, and reputation of the data) dimensions. Concerns with the trustworthiness of the social media and web scraped data are also raised (Tan et al., 2015). The efficiency of the physical flow of material can be determined by the

infrastructure quality (such as transportation system, ports, technology, etc.) (Bagchi et al., 2014). Therefore, for effective flow of information, it is vital to have advanced data infrastructures and best practices of data management using tools and techniques such as Hadoop, MapReduce and statistical process control.

c) *Lack of techniques and procedures:* Lack of quality data is not the only problem; there is an incapability of techniques to exploit the data deluge properly. For instance, in the case of demand forecasting techniques, significant attention is given solely to endogenous time-serious variables for demand forecasting, and there is a lack of consideration of exogenous variables and information sources (Meixell and Wu, 2001). This evidently has inferences to develop better data management capabilities and methodology for forecasting demand to enhance supply chain operations. Also, there is a difficulty in considering expert judgement as a covariate in forecasting models. Further, ordinal scales are used to measure opinions and experts' judgement. Whereas, in practice, the available modelling approach of independent and dependent variables are intended for continuous or natural data as input. As discussed in Blackburn et al. (2015), very few studies like Fildes et al. (2009) have reflected on using expert judgement to increase the accuracy of supply chain forecasting.

7. Implications for best practices of BDA in SCM

a) Firstly, for developing data generation capabilities, companies should start building their data infrastructure. Not all organisations have enormous data, there are many organisations, which do not generate big data by leveraging devices such as RFID and sensors. These organisations could start sensing their everyday business environment by deploying these novel technologies. On the other hand, organisations who are data rich should periodically audit their data sources such as RFID and sensors, to ensure no irrelevant data is being generated due to malfunctioning of these devices. Moreover, establishing metrics based on business rules, defining appropriate variables to measure, and deploying data reduction strategies could allow organisations to reduce the amount of data being generated at source (Rehman et al., 2016). The unstructured data (e.g. customer reviews) can be streamlined to generate less messy data. These approaches will certainly improve data quality and reduce infrastructure cost. When the generation of 'bad data' reduces, only good quality data will be created, thereby minimising data logistics efforts and storage costs.

b) Secondly, developing capabilities to integrate data across supply chains is associated with the practice of information sharing. There should be standard policies and principles established to standardise the process of information sharing (what to be and what not to be shared) with fellow supply chain members to gain mutual benefits and at the same time retain their individual competitive advantage. As we discussed before, cloud computing can be treated as a complementary resource which can aid information sharing maintaining security and access control.

c) Thirdly, for developing analytics capabilities, companies who have a low level of analytics capabilities could start from developing basic analytics (i.e. descriptive) and moving incrementally towards leveraging advanced analytics such as predictive and prescriptive. This paper suggests that predictive analytics is mostly used compared to other forms of analytics, but the analytics models once deployed should be validated periodically to ensure performance. Moreover, several factors could facilitate Big data practice; (i) Clearly defining the measurable outcomes, (ii) Effectively addressing the issues, (iii) Flexible model that enables stability and improves decision making, (iv) Utilising appropriate and most suitable data for the purpose (Radke and Tseng, 2015).

d) Fourthly, from the perspective of data-driven culture, the success of predictive analytics depends on the degree of integration into business processes and acceptance by management as a decision-making tool (Blackburn et al., 2015). Blackburn et al. (2015) suggested that inclusion of experts' opinion into analytics and better communication of analytics insights to management are the key ways of achieving integration and acceptance of analytics-driven decision making. Similarly, for successful implementation of Big Data analytics, it is critical to clearly define organisation-specific business requirements and performance measures, create a vision for utilising analytics using maturity models and defining roles based on analytical skills (Aho, 2015). The statistical understanding and critical thinking ability of decision makers can play a significant role in data-driven decision making (Markham et al., 2015; Radke and Tseng, 2015).

8. Conclusion and future research direction

In summary, BDA has the potential to outperform and transform traditional SCM practices. This study reviewed 82 academic papers on Big Data and SCM and 13 maturity models available in both academic and non-academic domains. In past, BDA has mostly been explored from the technological perspective to rationalise its economic benefits, but this

paper emphasised the necessity to delineate BDA capabilities in supply chain to extract value from big data. The structured approach used for literature review has revealed existing contributions of BDA and SCM research. Findings show a significant increase in the number of papers published in recent times. Social media based academic research has emerged as an important discipline in supply chain field. Findings suggest that BDA could be beneficial if organisations can develop the right capabilities to effectively use the big data.

Unlike prior research, the conceptualisation of BDA capabilities in this paper is holistic, and data-driven. First, it addresses the magnitude of understanding the provenance of big data in supply chain and optimising the data generation process. Second, it suggests the importance of integrating and standardising data from heterogeneous sources to offer more coherent data sets to analytics systems. Third, different types of analytics and the importance of assimilating the findings into the business process are addressed. In comparison with descriptive and predictive analytics, prescriptive analytics require a negligible amount of human intervention (Puget, 2015), which will revolutionise the decision-making process. Fourth, from value creation and users' perspective, the significant role of data visualisation and data-driven culture in increasing the flexibility and adaptability is discussed. Moreover, the efficacy of cloud computing and absorptive capacity in enhancing value from BDA investment is also explained.

This study makes significant contributions to both theory and practice. The conceptualisation of BDA capabilities would help academic researchers to embark on new empirical research in this domain. It contributes to the on-going debate of BDA in SCM context and supports a comprehensive understanding of this evolving technology from the systematic literature review and conceptualisation of key capabilities. A holistic conceptual framework of BDA capabilities is developed to describe the stages of BDA assimilation. Moreover, this paper will guide practitioners to realise their current state of BDA maturity and build a roadmap to develop BDA capabilities keeping in mind the potential challenges associated with assimilation process. Further, the paper also discussed the implications for best practices of BDA in supply chain.

While a systematic and structured literature review is conducted, it is worth recognising the concerns associated with this paper. It should be noted that the citation count information extracted from Scopus database differ from other databases such as WoS or Google scholar. The selection of journal papers was restricted by the use of specific

keywords, other keywords and databases could have resulted additional papers. Even though we relied on specific keywords and databases relevant for this research, efforts were made to validate, avoid errors and ensure that all relevant papers are included. Further, the synthesis of the conceptualisation does involve authors' interpretations of selected journal papers and their previous expertise in BDA domain. However, to ensure reliability the process of selecting journal papers, coding and conceptualisation were verified independently by peer researchers. Moreover, despite including maturity models from practice in the conceptualisation process, most of the papers reviewed are from academic sources. Only few challenges of BDA practice are discussed in this paper. Paper lacks in providing empirical evidence to support the conceptual framework, which could be addressed in future extension of this paper.

8.1 Direction for future research

i. The provenance of Big Data in supply chain context is not addressed in previous studies. Methods to optimise and deploy data generation infrastructure in supply chain network should be explored in future. Moreover, majority of papers discussed about extracting value from different types of data such as RFID data, transaction data and customer review data. But, more research should be required to develop methods and tools to analyse and integrate other forms of data including sensor, GPS, video and image data.

ii. With growing importance of data-driven decision making in supply chain management, poor data quality would negatively impact supply chain performance. Data quality should be monitored, measured and controlled. In future, a systematic investigation of various techniques used to tackle data quality problem in the context of inter-organisational supply chain networks should be investigated.

iii. In-memory analytics is advantageous for SCM as it enables real-time decision-making. While previous case studies (Piller and Hagedorn, 2011; vom Brocke et al., 2014) explained the potential benefits of in-memory analytics. There is a need for comprehensive investigation of the real-world use case of in-memory analytics in SCM.

iv. BDA influences a paradigm shift from heuristics to data-driven decision making. The reasoning and assimilation of data-driven insights into the process of decision-making and the behavioral issues associated with it have to be investigated. Moreover, the

phenomena of restriction to change from heuristics to data-driven decision making should be explored in various organisational contexts.

v. In recent times, papers have begun to focus on understanding the empirical relationship between the use of advanced analytics and supply chain performance. However, very few studies have addressed the underlying mechanism through which BDA can be utilised to support decision-making and business performance. The positive impact of BDA on business performance is not certain and the influence of several organisational factors remains ambiguous (Oliveira et al., 2012).

vi. There are various examples for successful implementation of BA to optimise supply chain and improve performance (Chae et al., 2014a). Along with academic sources, there are various anecdotal evidence such as company case studies, whitepapers, etc. justifying the positive impact of BDA on firm performance. However, there are some evidences of failure as well. Oliveira et al. (2012) and Sherer (2005) discussed the case of Cisco's BA failure. However, some questions remain unanswered, such as, why and under what circumstances organisations BDA initiative fails? There is a need for in-depth case studies to understand the organisational context under which organisations achieve success or failure from BDA investment.

vii. BDA implementation is a change-intensive process and organisations are vulnerable to disruption. Systems design should keep its pace in addressing these changing needs to achieve competitive advantage. Empirical studies are required to describe and understand the mechanisms and dynamics of how BDA implementation influences change and redesign of the system.

viii. The proportion of SMEs that have utilised big data analytics is found to be negligible (SAS, 2013). The disparity in adoption of these innovative technologies would certainly hinder the growth of SMEs. Perhaps, on the capabilities framework proposed in this paper, SMEs would fall under the category of Data Poor and Information Poor. It is important to investigate the effect of BDA adoption on the extension of the digital divide between SMEs and large organisations.

References

Acito, F., Khatri, V., 2014. Business analytics: Why now and what next? Bus. Horiz. 57, 565–570. doi:10.1016/j.bushor.2014.06.001

Aho, A.M., 2015. Product Data Analytics Service model for Manufacturing company, Lecture Notes in Business Information Processing, Lecture Notes in Business Information Processing. Springer International Publishing, Cham. doi:10.1007/978-3-319-21009-4

Alfaro, L.A., Le, T.M.H., Choi, H.R., Cho, M.J., 2015. Deployment Model of Big Data for Port Logistics. Inf. 18, 7–16.

Assunção, M.D., Calheiros, R.N., Bianchi, S., Netto, M.A.S., Buyya, R., 2015. Big data computing and clouds: Trends and future directions. J. Parallel Distrib. Comput. 79–80, 3–15. doi:10.1016/j.jpdc.2014.08.003

Autry, C.W., Grawe, S.J., Daugherty, P.J., Richey, R.G., 2010. The effects of technological turbulence and breadth on supply chain technology acceptance and adoption. J. Oper. Manag. 28, 522–536. doi:10.1016/j.jom.2010.03.001

Bagchi, P., Lejeune, M.A., Alam, A., 2014. How supply competency affects FDI decisions: Some insights. Int. J. Prod. Econ. 147, 239–251. doi:10.1016/j.ijpe.2013.02.025

Barlow, M., 2013. Real-Time Big Data Analytics: Emerging Architecture, 1st ed, Journal of Chemical Information and Modeling. O'Reilly Media, Inc. doi:10.1017/CBO9781107415324.004

Beyer, M.A., Laney, D., 2012. The Importance of "Big Data": A Definition. Gart. Publ. i, 1–9.

Bhattacharjya, J., Ellison, A., Tripathi, S., 2016. An exploration of logistics-related customer service provision on Twitter: The case of e-retailers. Int. J. Phys. Distrib. Logist. Manag. 46, 659–680.

Blackburn, R., Lurz, K., Priese, B., Göb, R., Darkow, I.-L., 2015. A predictive analytics approach for demand forecasting in the process industry. Int. Trans. Oper. Res. 22, 407–428. doi:10.1111/itor.12122

Bock, S.S., Isik, F.F., 2015. A new two-dimensional performance measure in purchase order sizing. Int. J. Prod. Res. 53, 4951–4962. doi:10.1080/00207543.2015.1005769

Bonnes, K., 2014. Predictive Analytics for Supply Chains : a Systematic Literature Review. Proc. BPM Demo Sess. 2014 Co-located with 12th Int. Conf. Bus. Process Manag. (BPM 2014).

Boone, C.A., Hazen, B.T., Skipper, J.B., Overstreet, R.E., 2016. A framework for investigating optimization of service parts performance with big data. Ann. Oper. Res. doi:10.1007/s10479-016-2314-1

Brandau, A., Tolujevs, J., 2013. Modelling and Analysis of Logistical State Data. Transp. Telecommun. 14, 102–115. doi:10.2478/ttj-2013-0009

Brooks, P., El-Gayar, O., Sarnikar, S., 2015. A framework for developing a domain specific business intelligence maturity model: Application to healthcare. Int. J. Inf. Manage. 35, 337–345. doi:10.1016/j.ijinfomgt.2015.01.011

Bruque Cámara, S., Moyano Fuentes, J., Maqueira Marín, J.M., 2015. Cloud computing, Web 2.0, and operational performance. Int. J. Logist. Manag. 26, 426–458. doi:10.1108/IJLM-07-2013-0085

Cao, G., Duan, Y., Li, G., 2015. Linking Business Analytics to Decision Making Effectiveness: A Path Model Analysis. IEEE Trans. Eng. Manag. 62, 384–395. doi:10.1109/TEM.2015.2441875

Chae, B. (Kevin), 2015. Insights from Hashtag #SupplyChain and Twitter Analytics: Considering Twitter and Twitter Data for Supply Chain Practice and Research. Int. J. Prod. Econ. 165, 247–259. doi:10.1016/j.ijpe.2014.12.037

Chae, B. (Kevin), Olson, D.L., 2013. Business analytics for supply chain: A dynamic-capabilities framework. Int. J. Inf. Technol. Decis. Mak. 12, 9–26. doi:10.1142/S0219622013500016

Chae, B. (Kevin), Yang, C., Olson, D., 2014a. The impact of supply chain analytics on operational performance: a resource-based view. Int. J. Prod. Res. 52, 4695–4710. doi:10.1080/00207543.2013.861616

Chae, B. (Kevin), Yang, C., Olson, D., Sheu, C., 2014b. The impact of advanced analytics and data accuracy on operational performance: A contingent resource based theory (RBT) perspective. Decis. Support Syst. 59, 119–126. doi:10.1016/j.dss.2013.10.012

Chan, H.K., Lacka, E., Yee, R.W.Y., Lim, M.K., 2015. The role of social media data in operations and production management. Int. J. Prod. Res. 7543, 1–10. doi:10.1080/00207543.2015.1053998

Chen, C.P.L., Zhang, C.Y., 2014. Data-intensive applications, challenges, techniques and technologies: A survey on Big Data. Inf. Sci. (Ny). 275, 314–347. doi:10.1016/j.ins.2014.01.015

Chen, D.Q., Preston, D.S., Swink, M., 2015. How the Use of Big Data Analytics Affects Value Creation in Supply Chain Management. J. Manag. Inf. Syst. 32, 4–39. doi:10.1080/07421222.2015.1138364

Chen, H., Chiang, R.H.L., Storey, V.C., 2012. Business intelligence and analytics: From big data to big impact. MIS Q. Manag. Inf. Syst. 36, 1165–1188.

Cohen, W., Levinthal, D., 1990. Absorptive Capacity : A New Perspective on Learning and Innovation. Adm. Sci. Q. 35, 128–152. doi:10.2307/2393553

Cosic, R., Shanks, G., Maynard, S., 2015. A business analytics capability framework. Australas. J. Inf. Syst. 19, 5–19.

Cox, M., Ellsworth, D., 1997. Application-controlled demand paging for out-of-core visualization. Proceedings. Vis. '97 (Cat. No. 97CB36155) 235–244,. doi:10.1109/VISUAL.1997.663888

Davenport, T.H., 2006. Competing on analytics [WWW Document]. Harv. Bus. Rev. URL https://hbr.org/2006/01/competing-on-analytics (accessed 2.2.16).

Davenport, T.H., Harris, J.G., De Long, D.W., Jacobson, A.L., 2001. Data to Knowledge to Results: BUILDING AN ANALYTIC CAPABILITY. Calif. Manage. Rev. 43, 117–138. doi:10.2307/41166078

Delen, D., Erraguntla, M., Mayer, R.J., Wu, C.N., 2011. Better management of blood supply-chain with GIS-based analytics. Ann. Oper. Res. 185, 181–193. doi:10.1007/s10479-009-0616-2

Dobrzykowski, D.D., Leuschner, R., Hong, P.C., Roh, J.J., 2015. Examining Absorptive Capacity in Supply Chains: Linking Responsive Strategy and Firm Performance. J. Supply Chain Manag. 51, 3–28. doi:10.1111/jscm.12085

Donovan, P.O., Leahy, K., Bruton, K., Sullivan, D.T.J.O., 2015. Big data in manufacturing : a systematic mapping study. J. Big Data. doi:10.1186/s40537-015-0028-x

Dutta, D., Bose, I., 2015. Managing a big data project: The case of Ramco Cements limited. Int. J. Prod. Econ. 165, 293–306. doi:10.1016/j.ijpe.2014.12.032

Ebner, K., Buhnen, T., Urbach, N., 2014. Think big with big data: Identifying suitable big data strategies in corporate environments. Proc. Annu. Hawaii Int. Conf. Syst. Sci. 3748–3757. doi:10.1109/HICSS.2014.466

Fahimnia, B., Sarkis, J., Davarzani, H., 2015. Green supply chain management: A review and bibliometric analysis. Int. J. Prod. Econ. 162, 101–114. doi:10.1016/j.ijpe.2015.01.003

Fawcett, S.E.., 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). J. Bus. Logist. 35, 157–164. doi:10.1111/jbl.12058

Few, S., 2009. Now You See It: Simple Visualization Techniques for Quantitative Analysis, Distribution. Analytics Press. doi:10.1016/S0262-4079(06)60844-0

Fichman, R.G., Kemerer, C.F., 1997. The assimilation of software process innovations : an organizational learning perspective. Manage. Sci. 43, p1345-1363. doi:10.1287/mnsc.43.10.1345

Fildes, R., Goodwin, P., Lawrence, M., Nikolopoulos, K., 2009. Effective forecasting and judgmental adjustments: an empirical evaluation and strategies for improvement in supply-chain planning. Int. J. Forecast. 25, 3–23. doi:10.1016/j.ijforecast.2008.11.010

Wamba, S. F., Akter, S., 2015. Big data analytics for supply chain management: A literature review and research agenda, Lecture Notes in Business Information Processing, Lecture Notes in Business Information Processing. Springer International Publishing, Cham. doi:10.1007/978-3-319-24626-0

Gandomi, A., Haider, M., 2015. Beyond the hype: Big data concepts, methods, and analytics. Int. J. Inf. Manage. 35, 137–144. doi:10.1016/j.ijinfomgt.2014.10.007

Gao, D., Xu, Z., Ruan, Y.Z., Lu, H., 2016. From a systematic literature review to integrated definition for sustainable supply chain innovation (SSCI). J. Clean. Prod. doi:10.1016/j.jclepro.2016.11.153

Gartner, 2014. Gartner Predicts Business Intelligence and Analytics Will Remain Top Focus for CIOs Through 2017 [WWW Document]. Press release. URL http://www.gartner.com/newsroom/id/2637615 (accessed 8.18.16).

Ge, X., Jackson, J., 2014. The Big Data Application Strategy for Cost Reduction in Automotive Industry. SAE Int. J. Commer. Veh. 7, 588–598. doi:10.4271/2014-01-2410

Gudfinnsson, K., Strand, M., Berndtsson, M., 2015. Analyzing business intelligence maturity.

Transportation Research Part E: Logistics and Transportation Review 2018; 114: 416-436.

J. Decis. Syst. 24, 37–54. doi:10.1080/12460125.2015.994287

Gunasekaran, A., Papadopoulos, T., Dubey, R., Wamba, S.F., Childe, S.J., Hazen, B., Akter, S., 2017. Big data and predictive analytics for supply chain and organizational performance. J. Bus. Res. 70, 308–317. doi:10.1016/j.jbusres.2016.08.004

Hahn, G.J., Packowski, J., 2015. A Perspective on Applications of In-memory Analytics in Supply Chain Management. Decis. Support Syst. 76, 45–52. doi:10.1016/j.dss.2015.01.003

Halper, F., Krishnan, K., 2014. TDWi Big Data Maturity Model Guide. TDWI Resarch 2013–2014, 1–20.

Hazen, B.T., Boone, C.A., Ezell, J.D., 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. Int. J. Prod. Econ. 154, 72–80. doi:10.1016/j.ijpe.2014.04.018

Hazen, B.T., Skipper, J.B., Ezell, J.D., Boone, C.A., 2016. Big Data and Predictive Analytics for Supply Chain Sustainability: A Theory-driven Research Agenda. Comput. Ind. Eng. doi:10.1016/j.cie.2016.06.030

Hofmann, E., 2015. Big data and supply chain decisions: the impact of volume, variety and velocity properties on the bullwhip effect. Int. J. Prod. Res. 7543, 1–19. doi:10.1080/00207543.2015.1061222

Hu, H., Wen, Y., Chua, T.S., Li, X., 2014. Toward Scalable Systems for Big Data Analytics: A Technology Tutorial. Access, IEEE 2, 652–687. doi:10.1109/ACCESS.2014.2332453

Hurwitz, J., Nugent, A., Halper, F., Kaufman, M., 2013. Big Data For Dummies. John Wiley & Sons, Hoboken, New Jersey. doi:10.1007/s13398-014-0173-7.2

IDC, 2013. CSC Big data maturity tool. Orig. Dev. by IDC, CSC Off. website. URL http://csc.bigdatamaturity.com/ (accessed 8.5.16).

Jamehshooran, B.G., Shaharoun, A.M., Norehan, H.H., 2015a. The moderating effect of web service on the relationship between business analytics and supply chain performance. J. Theor. Appl. Inf. Technol. 76, 97–108.

Jamehshooran, B.G., Shaharoun, A.M., Norehan, H.H., 2015b. Assessing Supply Chain Performance through Applying the SCOR Model. Int. J. Supply Chain Manag. 4, 1–11.

Janssen, M., Estevez, E., Janowski, T., 2014. Interoperability in big, open, and linked data-organizational maturity, capabilities, and data portfolios. Computer (Long. Beach. Calif). 47, 44–49. doi:10.1109/MC.2014.290

Jin, Y., Ji, S., 2013. Partner Choice of Supply Chain Based on 3d Printing and Big Data. Inf. Technol. J. 12, 6822–6826. doi:10.3923/itj.2013.6822.6826

Kabir, N., Carayannis, E., 2013. Big Data , Tacit Knowledge and Organizational Competitiveness. J. Intell. Stud. Bus. 3, 220–228.

Kang, Y., Park, I., Rhee, J., Lee, Y., 2016. MongoDB-based Repository Design for IoT-generated RFID / Sensor Big Data. IEEE Sens. J. 16, 485–497. doi:10.1109/JSEN.2015.2483499

Kirk, A., 2012. Data Visualization : a successful design process. Packt Publishing Ltd.

Kiron, D., Shockley, R., 2011. Creating business value with analytics. MIT Sloan Manag. Rev. 53, 57–63.

Knowledgent, 2014. Big data maturity assessment. Knowledgent Gr. Inc. Off. website. URL https://bigdatamaturity.knowledgent.com (accessed 8.5.16).

Koh, S.C.L., Genovese, A., Acquaye, A.A., Barratt, P., Kuylenstierna, J., Gibbs, D., 2013. Decarbonising product supply chains : design and development of an integrated evidence- based decision support system – the supply chain environmental analysis tool (SCEnAT). Int. J. Prod. Res. 51, 2092–2109. doi:10.1080/00207543.2012.705042

Kraus, C., Valverde, R., 2014. A data warehouse design for the detection of fraud in the supply chain by using the benford's law. Am. J. Appl. Sci. 11, 1507–1518. doi:10.3844/ajassp.2014.1507.1518

Laursen, G.H.N., Thorlund, J., 2010. Business Analytics for Managers: Taking Business Intelligence Beyond Reporting, Wiley and SAS business series. John Wiley & Sons, Inc, New Jersey.

Lavalle, S., Hopkins, M.S., Lesser, E., Shockley, R., Kruschwitz, N., 2010. Analytics : The New Path to Value. MIT Sloan Manag. Rev. 1–24.

Lee, Y.W., Strong, D.M., Kahn, B.K., Wang, R.Y., 2002. AIMQ: a methodology for information quality assessment. Inf. Manag. 40, 133–146. doi:10.1016/S0378-7206(02)00043-5

Li, L., Chi, T., Hao, T., Yu, T., 2016. Customer demand analysis of the electronic commerce supply chain using Big Data. Ann. Oper. Res. doi:10.1007/s10479-016-2342-x

Lin, H.-F., 2014. Understanding the determinants of electronic supply chain management system adoption: Using the technology–organization–environment framework. Technol. Forecast. Soc. Change 86, 80–92. doi:10.1016/j.techfore.2013.09.001

Malhotra, A., Gosain, S., El Sawy, O.A., 2005. Absorptive capacity configurations in supply chains: Gearing for parter-enabled market knowledge creation. MIS Q. 29, 145–187. doi:10.2307/25148671

Manyika, J., Chui, M., Brown, B., Bughin, J., Dobbs, R., Roxburgh, C., Hung, A., Byers, 2011. Big data: The next frontier for innovation, competition, and productivity, McKinsey Global Institute. doi:10.1080/01443610903114527

Markham, S.K., Kowolenko, M., Michaelis, T.L., 2015. Unstructured Text Analytics to Support New Product Development Decisions. Res. Technol. Manag. 58, 30–39. doi:10.5437/08956308X5802291

Marr, B., 2014. Big data: 25 amazing need-to-know facts. URL http://smartdatacollective.com/bernardmarr/277731/big-data-25-facts-everyone-needs-know (accessed 3.12.16).

Mayring, P., 2003. Qualitative Content Analysis, eighth ed. ed. Beltz Verlag, Weinheim, Germany.

Meixell, M., Wu, S., 2001. Scenario analysis of demand in a technology market using leading indicators. Semicond. Manuf. IEEE … 14, 65–75.

Meredith, R., Remington, S., O'Donnell, P., Sharma, N., 2012. Organisational transformation

through Business Intelligence: theory, the vendor perspective and a research agenda. J. Decis. Syst. 21, 187–201. doi:10.1080/12460125.2012.731218

Mishra, D., Gunasekaran, A., Papadopoulos, T., Childe, S.J., Oper Res, A., 2016. Big Data and supply chain management: a review and bibliometric analysis. Ann. Oper. Res. doi:10.1007/s10479-016-2236-y

Mishra, N., Singh, A., 2016. Use of twitter data for waste minimisation in beef supply chain. Ann. Oper. Res. doi:10.1007/s10479-016-2303-4

Mortenson, M.J.., Doherty, N.F., Robinson, S., 2015. Operational research from Taylorism to Terabytes: A research agenda for the analytic sage. Eur. J. Oper. Res. 241, 583–595. doi:doi:10.1016/j.ejor.2014.08.029

Neaga, I., Liu, S., Xu, L., Chen, H., Hao, Y., 2015. Cloud enabled big data business platform for logistics services: A research and development agenda, in: Delibašić, B., Hernández, E.J., Papathanasiou, J., Dargam, F., Zaraté, P., Ribeiro, R., Liu, S., Linden, I. (Eds.), Lecture Notes in Business Information Processing. Springer International Publishing, Cham, pp. 22–33. doi:10.1007/978-3-319-18533-0_3

Nott, C., 2014. Big Data & Analytics Maturity Model. IBM Off. website. URL http://www.ibmbigdatahub.com/blog/big-data-analytics-maturity-model (accessed 8.5.16).

Oliveira, M.P.V. de, McCormack, K., Trkman, P., 2012. Business analytics in supply chains – The contingent effect of business process maturity. Expert Syst. Appl. 39, 5488–5498. doi:10.1016/j.eswa.2011.11.073

Oxford Dictionary, 2016. In: Oxford Dictionary. URL http://www.oxforddictionaries.com/definition/english/data (accessed 5.12.16).

Papadopoulos, T., Gunasekaran, A., Dubey, R., Altay, N., Childe, S.J., Fosso- Wamba, S., 2016. The role of Big Data in explaining disaster resilience in supply chains for sustainability. J. Clean. Prod. doi:10.1016/j.jclepro.2016.03.059

Park, H., Bellamy, M.A., Basole, R.C., 2016. Visual analytics for supply network management: System design and evaluation. Decis. Support Syst. 91, 89–102. doi:http://dx.doi.org/10.1016/j.dss.2016.08.003

Piller, G., Hagedorn, J., 2011. Business Benefits And Application Capabilities Enabled By In-Memory Data Management. Proc. des Work. Innov. Unternehmensanwendungen mit In-Memory Data Manag. 45–56.

Popovič, A., Hackney, R., Coelho, P.S., Jaklič, J., 2012. Towards business intelligence systems success: Effects of maturity and culture on analytical decision making. Decis. Support Syst. 54, 729–739. doi:10.1016/j.dss.2012.08.017

Puget, J.F., 2015. The Analytics Maturity Model. IBM Off. website. URL https://www.ibm.com/developerworks/community/blogs/jfp/entry/the_analytics_maturity_model?lang=en (accessed 8.4.16).

Radcliffe, J., 2014. Leverage a Big Data Maturity Model to Build Your Big Data Roadmap. Radcliffe Advis. Serv. Ltd.

Radke, A.M., Tseng, M.M., 2015. Design Considerations for Building Distributed Supply Chain Management Systems Based on Cloud Computing. J. Manuf. Sci. Eng. 137,

40906. doi:10.1115/1.4030736

Rehman, M.H. ur, Chang, V., Batool, A., Wah, T.Y., 2016. Big data reduction framework for value creation in sustainable enterprises. Int. J. Inf. Manage. 36, 917–928. doi:http://dx.doi.org/10.1016/j.ijinfomgt.2016.05.013

Richey Jr, R.G., Morgan, T.R., Lindsey-Hall, K., Adams, F.G., 2016. A global exploration of Big Data in the supply chain. Int. J. Phys. Distrib. Logist. Manag. 46, 710–739.

Roberts, N., Galluch, P.S., Dinger, M., Grover, V., 2012. Absorptive Capacity and Information Systems Research: Review, Synthesis, and Directions for Future Research. MIS Q. 36, 625–648.

Rozados, V.I., Tjahjono, B., 2014. Big Data Analytics in Supply Chain Management : Trends and Related Research, in: 6th International Conference on Operations and Supply Chain Management, Bali.

Sahay, B.S., Ranjan, J., 2008. Real time business intelligence in supply chain analytics. Inf. Manag. Comput. Secur. 16, 28–48. doi:10.1108/09685220810862733

Sanders, N.R., 2016. How to Use Big Data to Drive Your Supply Chain. Calif. Manage. Rev. 58, 26–48. doi:10.1525/cmr.2016.58.3.26

Sangari, M.S., Razmi, J., 2015. Business intelligence competence, agile capabilities, and agile performance in supply chain: An empirical study. Int. J. Logist. Manag. 26, 356–380. doi:10.1108/IJLM-01-2013-0012

SAS, 2013. Big Data Analytics Adoption and Employment Trends, 2012-2017. London.

Saunders, M., Lewis, P., Thornhill, A., 2016. Research Methods For Business Students, 7th ed, Pearson Education UK. Pearson Education Limited. doi:10.1017/CBO9781107415324.004

Schoenherr, T., Speier-Pero, C., 2015. Data Science, Predictive Analytics, and Big Data in Supply Chain Management: Current State and Future Potential. J. Bus. Logist. 36, 120–132. doi:10.1111/jbl.12082

Seah, M., Hsieh, M.H., Weng, P.-D., 2010. A case analysis of Savecom: The role of indigenous leadership in implementing a business intelligence system. Int. J. Inf. Manage. 30, 368–373. doi:10.1016/j.ijinfomgt.2010.04.002

See-To, E.W.K., Ngai, E.W.T., 2016. Customer reviews for demand distribution and sales nowcasting: a big data approach. Ann. Oper. Res. 1–17. doi:10.1007/s10479-016-2296-z

Seuring, S., Müller, M., 2008. From a literature review to a conceptual framework for sustainable supply chain management. J. Clean. Prod. 16, 1699–1710. doi:10.1016/j.jclepro.2008.04.020

Sherer, S.A., 2005. From supply chain management to value network advocacy: implications for e-supply chains. Supply Chain Manag. An Int. J. 10, 77–83. doi:10.1108/13598540510589151

Shim, J.P., Warkentin, M., Courtney, J.F., Power, D.J., Sharda, R., Carlsson, C., 2002. Past, present, and future of decision support technology. Decis. Support Syst. 33, 111–126. doi:10.1016/S0167-9236(01)00139-7

Simpson, J.A., Weiner, E.S.C., 1989. The Oxford English Dictionary, Oxford University

Press, Oxford, UK.

Singh, A., Mishra, N., Ali, S.I., Shukla, N., Shankar, R., 2015. Cloud computing technology: Reducing carbon footprint in beef supply chain. Int. J. Prod. Econ. 164, 462–471. doi:10.1016/j.ijpe.2014.09.019

Sonka, S., 2014. Big data and the ag sector: More than lots of numbers. Int. Food Agribus. Manag. Rev. 17, 1–20.

Souza, G.C., 2014. Supply chain analytics. Bus. Horiz. 57, 595–605. doi:10.1016/j.bushor.2014.06.004

Spruit, M., Sacu, C., 2015. DWCMM: The Data Warehouse Capability Maturity Model. J. Univers. Comput. Sci. 21, 1508–1534.

Stefanovic, N., 2015. Collaborative Predictive Business Intelligence Model for Spare Parts Inventory Replenishment. Comput. Sci. Inf. Syst. 12, 911. doi:10.2298/CSIS141101034S

Stefanovic, N., 2014. Proactive Supply Chain Performance Management with Predictive Analytics. Sci. World J. 1–17.

Sulaiman, H., Cob, Z.C., Ali, N., 2015. Big Data Maturity Model for Malaysian Zakat Institutions to Embark on Big Data Initiatives. 4th Int. Conf. Softw. Eng. Comput. Syst. 61–66.

Tachizawa, E.M., Alvarez-Gil, M.J., Montes-Sancho, M.J., 2015. How "smart cities" will change supply chain management. Supply Chain Manag. An Int. J. 20, 237–248. doi:10.1108/SCM-03-2014-0108

Tan, K.H., Zhan, Y., Ji, G., Ye, F., Chang, C., 2015. Harvesting big data to enhance supply chain innovation capabilities: An analytic infrastructure based on deduction graph. Int. J. Prod. Econ. 165, 223–233. doi:10.1016/j.ijpe.2014.12.034

The Economist, 2010. The data deluge.

Tien, J.M., 2015. Internet of connected ServGoods: Considerations, consequences and concerns. J. Syst. Sci. Syst. Eng. 24, 130–167. doi:10.1007/s11518-015-5273-1

Tien, J.M., 2012. The next industrial revolution: Integrated services and goods. J. Syst. Sci. Syst. Eng. 21, 257–296. doi:10.1007/s11518-012-5194-1

Trkman, P., Ladeira, M.B., Valadares De Oliveira, M.P., McCormack, K., 2012. Business Analytics, Process Maturity and Supply Chain Performance. Bus. Process Manag. Work. Pt I 99, 111–122. doi:10.1007/978-3-642-28108-2_10

Trkman, P., McCormack, K., de Oliveira, M.P.V., Ladeira, M.B., 2010. The impact of business analytics on supply chain performance. Decis. Support Syst. 49, 318–327. doi:10.1016/j.dss.2010.03.007

van der Spoel, S., Amrit, C., van Hillegersberg, J., 2015. Predictive analytics for truck arrival time estimation: a field study at a European distribution center. Int. J. Prod. Res. 1–17. doi:10.1080/00207543.2015.1064183

Vera-Baquero, A., Colomo Palacios, R., Stantchev, V., Molloy, O., 2015. Leveraging big-data for business process analytics. Learn. Organ. 22, 215–228. doi:10.1108/TLO-05-2014-0023

vom Brocke, J., Debortoli, S., Müller, O., Reuter, N., 2014. How in-memory technology can create business value: Insights from the hilti case. Commun. Assoc. Inf. Syst. 34, 151–168.

Waller, M.A., Fawcett, S.E., 2013. Data Science, Predictive Analytics, and Big Data: A Revolution That Will Transform Supply Chain Design and Management. J. Bus. Logist. 34, 77–84. doi:10.1111/jbl.12010

Wamba, S.F., Akter, S., Edwards, A., Chopin, G., Gnanzou, D., 2015. How "big data" can make big impact: Findings from a systematic review and a longitudinal case study. Int. J. Prod. Econ. 165, 234–246. doi:10.1016/j.ijpe.2014.12.031

Wang, G., Gunasekaran, A., Ngai, E.W.T., Papadopoulos, T., 2016a. Big data analytics in logistics and supply chain management: Certain investigations for research and applications. Int. J. Prod. Econ. 176, 98–110. doi:10.1016/j.ijpe.2016.03.014

Wang, R.Y., 1998. A Product Perspective on Total Data Quality Management. Commun. ACM 41, 58–65. doi:10.1145/269012.269022

Wang, R.Y., Storey, V.C., Firth, C.P., 1995. A Framework for Analysis of Data Quality Research. IEEE Trans. Knowl. Data Eng. 7.

Wang, R.Y., Strong, D.M., 1996. Beyond Accuracy: What Data Quality Means to Data Consumers. J. Manag. Inf. Syst. 12, 5–33.

Wang, W., Liu, L., Feng, Y., Wang, T., 2014. Innovation with IS usage: individual absorptive capacity as a mediator. Ind. Manag. Data Syst. 144, 1110–1130. doi:10.1108/IMDS-05-2014-0160

Wang, Y., Kung, L., Byrd, T.A., 2016b. Big data analytics: Understanding its capabilities and potential benefits for healthcare organizations. Technol. Forecast. Soc. Change. doi:10.1016/j.techfore.2015.12.019

Xiong, G., Nyberg, T.R., Dong, X., Shang, X., 2015. Intelligent technologies and systems of material management, in: Kahraman, C., Çevik Onar, S. (Eds.), Intelligent Systems Reference Library, Intelligent Systems Reference Library. Springer International Publishing, Cham, pp. 295–330. doi:10.1007/978-3-319-17906-3

Yesudas, M., Menon, G., Ramamurthy, V., 2014. Intelligent operational dashboards for smarter commerce using big data. IBM J. Res. Dev. 58, 13:1-13:10. doi:10.1147/JRD.2014.2346131

Zhang, Q., Huang, T., Zhu, Y., Qiu, M., 2013. A Case Study of Sensor Data Collection and Analysis in Smart City: Provenance in Smart Food Supply Chain. Int. J. Distrib. Sens. Networks 1–12. doi:10.1155/2013/382132

Zhong, R.Y., Huang, G.Q., Lan, S., Dai, Q.Y., Chen, X., Zhang, T., 2015. A Big Data Approach For Logistics Trajectory Discovery From Rfid-Enabled Production Data. Int. J. Prod. Econ. 165, 260–272. doi:10.1016/j.ijpe.2015.02.014

Zhong, R.Y., Lan, S., Xu, C., Dai, Q., Huang, G.Q., 2016a. Visualization of RFID-enabled shopfloor logistics Big Data in Cloud Manufacturing. Int. J. Adv. Manuf. Technol. 84, 5–16. doi:10.1007/s00170-015-7702-1

Zhong, R.Y., Newman, S.T., Huang, G.Q., Lan, S., 2016b. Big Data for supply chain management in the service and manufacturing sectors: Challenges, opportunities, and

future perspectives. Comput. Ind. Eng. doi:10.1016/j.cie.2016.07.013