Exploring big data analytics capabilities for supply chain: a systematic literature review

Deepak Arunachalam
Sheffield University Management School, Conduit Road, Sheffield, S10 1FL.
email: darunachalam1@sheffield.ac.uk

Niraj Kumar
Sheffield University Management School, Conduit Road, Sheffield, S10 1FL.

John Paul Kawalek
Sheffield University Management School, Conduit Road, Sheffield, S10 1FL.

Nagesh Shukla
Faculty of Engineering and Information Sciences, University of Wollongong NSW 2522 Australia.

Abstract

In the era of Big Data, many organisations have successfully leveraged BDA capabilities to extract value from data to improve their performance. Development of BDA capabilities for supply chain management is crucial for business growth. This paper aims to provide a systematic literature review of BDA capabilities in context of supply chain and develop the capabilities maturity model for SCM. A conceptual framework to link BDA capabilities with different stages of generating and assimilating the added value is proposed in this paper. This paper contributes in theorizing BDA capabilities, and provides propositions for future research.

Keywords: Big Data Analytics, Supply chain, Systematic literature review.

Introduction

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. Adoption of BDA technologies could improve organisation capabilities in today's rapidly changing dynamic market environment (Meredith et al. 2012). To effectively deal with the diffusion of BDA technologies into the supply chain, organisational and behavioural issues associated with BDA adoption and practice has to be addressed. However, 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 reviewed existing literature and finds very few empirical studies discussing the impact of BDA capabilities on SCM. In addition, academic research 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. Finally, a conceptual model is developed and propositions are also discussed.

Research approach

The literature review approach proposed by Mayring (2003) is adopted in this paper. A 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. Further, 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, a deductive approach is used as new codes have to be devised adaptably as it emerges from the data (Saunders et al. 2016, p.582).

Several terms related to BDA 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, various keywords related to SCM are also identified. Different combinations of these 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. Moreover, 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. Initial shortlisting has produced 619 papers. After removing the duplication and verifying it in Endnote software, the full text of remaining papers is 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. Further, 13 maturity models from academic and industry sources are also shortlisted for the review process. 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 brief findings from the descriptive analysis is discussed in the next section.

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

Findings of literature review

Findings suggests that there is an increasing trend in terms of a number of papers published in the field of BDA in supply chain. Among 82 papers selected, most of them were published in last four years, which indicates rising importance among researchers to investigate the phenomenon of BDA in supply chain context. The top 10 contributing authors were extracted using BibExcel tool. It was found out 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 these criteria, followed by Chae and Gunasekaran. Further, the prominent theories used to explain the phenomenon of BDA adoption and practice are also reviwed. 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. The research methods used by selected papers are also reviewed. Most of the studies have used conceptual approach, followed by case studies and Survey-based research. 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. Among the selected papers, Hu et.al. (2014), Trkman et al. (2010) and Waller and Fawcett (2013 b) have received the highest number of citations. Trkman et al. (2010) was the first to quantitatively investigate the impact of business analytics on supply chain performance, followed by several other researchers.

Conceptual model development

Using thematic analysis, the reviewed papers are further 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.

- *Data Generation (DG) capability* is the ability of organisations to seek, identify, create, and access data from heterogeneous data sources across organisational boundaries.
- *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.
- *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 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-driven culture* is an intangible resource that represents the beliefs, attitudes, and opinion of people towards data-driven decision-making.

Further, a conceptual model is developed to explain the role of BDA capabilities maturity and organisational absorptive capacity in enhancing organisational performance. Consequently, two propositions are discussed in this paper based on the model shown in figure 1.

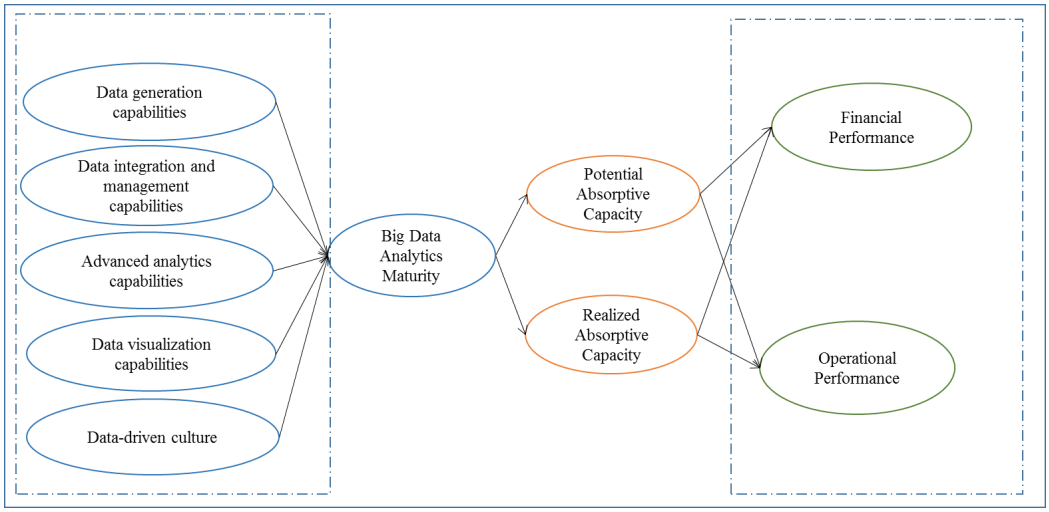

Figure 1: The conceptual model

Proposition 1: The level of BDA maturity is positively associated with the degree of data quality and the level of organisational performance.

The 5 key dimensions of BDA capabilities discussed earlier collectively constitutes to the BDA maturity of organisations. In the 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. 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). Moreover, vast amount of supply chain data is distributed in heterogeneous sources and integration of these isolated data will be challenging (Stefanovic 2014). 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). Due to the dynamic nature of supply chains, real-time access, integration and scalability of data storage are the key capabilities to possess.

In addition, a 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 has found to be create economic value. 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. 2014). Besides, data visualisation capability is equally important compared to other BDA capabilities. Park et al. (2016) have developed 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. Finally, 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. According to Aho (2015, p.284) "The transformative potential of Big Data lies in treating data as an asset." Real-world case examples have suggested engagement of implementation team and top management support are significant for developing BDA capabilities (Wamba et al. 2015). 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).Hence, it can be argued possessing BDA capabilities would enhance the decision-making and organisation performance.

Proposition 2: The greater the level of BDA capabilities maturity, the more enhanced is the organisations 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. 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.

Conclusion

In summary, BDA has the potential to outperform and transform traditional SCM practices. This study reviewed 82 academic papers on Big Data and SCM 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. 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. The few propositions discussed in this paper will be addressed in future research. Therefore, 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. 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.

References

Aho, A.M. (2015). Product Data Analytics Service model for Manufacturing company. L. Uden, M. Heričko, & I.-H. Ting, eds. Cham: Springer International Publishing.

Bhattacharjya, J., Ellison, A. and Tripathi, S. (2016). An exploration of logistics-related customer service provision on Twitter: The case of e-retailers. International Journal of Physical Distribution & Logistics Management, 46(6/7), pp.659–680.

Blackburn, R. et al. (2015). A predictive analytics approach for demand forecasting in the process industry. International Transactions in Operational Research, 22(3), pp.407–428.

Brandau, A. and Tolujevs, J. (2013). Modelling and Analysis of Logistical State Data. Transport and Telecommunication, 14(2), pp.102–115.

Cao, G., Duan, Y. and Li, G. (2015). Linking Business Analytics to Decision Making Effectiveness: A Path Model Analysis. IEEE Transactions on Engineering Management, 62(3), pp.384–395.

Chae, B. (Kevin). (2015). Insights from Hashtag #SupplyChain and Twitter Analytics: Considering Twitter and Twitter Data for Supply Chain Practice and Research. International Journal of Production Economics, 165, pp.247–259.

Chae, B. (Kevin), Yang, C. and Olson, D. (2014). The impact of supply chain analytics on operational performance: a resource-based view. International Journal of Production Research, 52(16), pp.4695–4710.

Chan, H.K. et al. (2015). The role of social media data in operations and production management. International Journal of Production Research, 7543(July 2015), pp.1–10.

Chen, D.Q., Preston, D.S. and Swink, M. (2015). How the Use of Big Data Analytics Affects Value Creation in Supply Chain Management. Journal of Management Information Systems, 32(4), pp.4–39.

Cohen, W. and Levinthal, D. (1990). Absorptive Capacity : A New Perspective on Learning and Innovation. Administrative science quarterly, 35(1), pp.128–152.

Davenport, T.H. (2006). Competing on analytics. Harvard Business Review, (84), pp.98–107.

Delen, D. et al. (2011). Better management of blood supply-chain with GIS-based analytics. Annals of Operations Research, 185(1), pp.181–193.

Dobrzykowski, D.D. et al. (2015). Examining Absorptive Capacity in Supply Chains: Linking Responsive Strategy and Firm Performance. Journal of Supply Chain Management, 51(4), pp.3–28.

Fahimnia, B., Sarkis, J. and Davarzani, H. (2015). Green supply chain management: A review and bibliometric analysis. International Journal of Production Economics, 162, pp.101–114.

Gao, D. et al. (2016). From a systematic literature review to integrated definition for sustainable supply chain innovation (SSCI). Journal of Cleaner Production.

Hazen, B.T. et al. (2016). Big Data and Predictive Analytics for Supply Chain Sustainability: A Theory-driven Research Agenda. Computers & Industrial Engineering.

Hofmann, E. (2015). Big data and supply chain decisions: the impact of volume, variety and velocity properties on the bullwhip effect. International Journal of Production Research, 7543(December 2015), pp.1–19.

Li, L. et al. (2016). Customer demand analysis of the electronic commerce supply chain using Big Data. Annals of Operations Research.

Malhotra, A., Gosain, S. and El Sawy, O.A. (2005). Absorptive capacity configurations in supply chains: Gearing for parter-enabled market knowledge creation. MIS Quarterly, 29(1), pp.145–187.

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

Meredith, R. et al. (2012). Organisational transformation through Business Intelligence: theory, the vendor perspective and a research agenda. Journal of Decision Systems, 21(3), pp.187–201.

Mishra, N. and Singh, A. (2016). Use of twitter data for waste minimisation in beef supply chain. Annals of Operations Research.

Papadopoulos, T. et al. (2016). The role of Big Data in explaining disaster resilience in supply chains for sustainability. Journal of Cleaner Production.

Park, H., Bellamy, M.A. and Basole, R.C. (2016). Visual analytics for supply network management: System design and evaluation. Decision Support Systems, 91, pp.89–102.

Popovič, A. et al. (2012). Towards business intelligence systems success: Effects of maturity and culture on analytical decision making. Decision Support Systems, 54(1), pp.729–739.

Roberts, N. et al. (2012). Absorptive Capacity and Information Systems Research: Review, Synthesis, and Directions for Future Research. MIS Quarterly, 36(2), pp.625–648.

Sanders, N.R. (2016). How to Use Big Data to Drive Your Supply Chain. California Management Review, 58(3), pp.26–48.

Saunders, M., Lewis, P. and Thornhill, A. (2016). Research Methods For Business Students. 7th ed. Pearson Education Limited.

See-To, E.W.K. and Ngai, E.W.T. (2016). Customer reviews for demand distribution and sales nowcasting: a big data approach. Annals of Operations Research, pp.1–17.

Seuring, S. and Müller, M. (2008). From a literature review to a conceptual framework for sustainable supply chain management. Journal of Cleaner Production, 16(15), pp.1699–1710.

van der Spoel, S., Amrit, C. and van Hillegersberg, J. (2015). Predictive analytics for truck arrival time estimation: a field study at a European distribution center. International Journal of Production Research, (May), pp.1–17.

Stefanovic, N. (2014). Proactive Supply Chain Performance Management with Predictive Analytics. The Scientific World Journal, pp.1–17.

Tan, K.H. et al. (2015). Harvesting big data to enhance supply chain innovation capabilities: An analytic infrastructure based on deduction graph. International Journal of Production Economics, 165, pp.223–233.

Trkman, P. et al. (2010). The impact of business analytics on supply chain performance. Decision Support Systems, 49(3), pp.318–327.

Wamba, S.F. et al. (2015). How 'big data' can make big impact: Findings from a systematic review and a longitudinal case study. International Journal of Production Economics, 165, pp.234–246.

Xiong, G. et al. (2015). Intelligent technologies and systems of material management. In C. Kahraman & S. Çevik Onar, eds. Intelligent Systems Reference Library. Intelligent Systems Reference Library. Cham: Springer International Publishing, pp. 295–330.

Zhang, Q. et al. (2013). A Case Study of Sensor Data Collection and Analysis in Smart City: Provenance in Smart Food Supply Chain. International Journal of Distributed Sensor Networks, pp.1–12.

Zhong, R.Y. et al. (2015). A Big Data Approach For Logistics Trajectory Discovery From Rfid-Enabled Production Data. International Journal of Production Economics, 165, pp.260–272.

Zhong, R.Y. et al. (2016). Visualization of RFID-enabled shopfloor logistics Big Data in Cloud Manufacturing. International Journal of Advanced Manufacturing Technology, 84(1), pp.5–16.