# Big Data Analytics for Supply Chain Management: a review of empirical studies

*Julian Klein*
*University of Münster*

*Raquel Mello (raquel.mello@ercis.uni-muenster.de)*
*University of Münster*

*Ann-Kristin Cordes*
*University of Münster*

*Bernd Hellingrath*
*University of Münster*

*Roberto Antonio Martins*
*Federal University of São Carlos*

## Abstract

Big data analytics (BDA) has emerged as a relevant topic in the supply chain management (SCM) area. Many studies in SCM show the benefits and challenges induced by BDA implementation. However, there is a lack of empirical evidence besides the great consent among researchers on the positive impact. This paper aims to fill this gap conducting a systematic literature review to find the academic literature focusing on empirical studies. The empirical evidence from those papers is discussed regarding benefits and challenges comparing them to the theory. Most of benefits and challenges from empirical studies corroborate the theory.

**Keywords:** Supply Chain Management, Big Data Analytics, Systematic Literature Review

### Introduction

Big data has gained increasing interest both in industry and academic research (Mishra et al., 2017). The 5V (volume, variety, velocity, value, and veracity) are the elements of a general description of big data. The big data definition is a holistic approach to generate, gather, store, manage, process, and analyze large datasets with high velocity and variety (Demchenko et al., 2013; Sanders, 2014; Wamba et al., 2015; Gandomi and Haider, 2015). Big data analytics (BDA) is the application of advanced analytical techniques in big data with an enormous potential to impact positively on business and academic scenarios (Russom, 2011). The BDA application can provide valuable insight

for decision makers, leading to a multitude of benefits that will lead towards a competitive advantage (Wamba et al., 2015).

Similarly, researchers from the field of supply chain management (SCM) have gained interest in BDA. Brinch et al. (2017) argue that big data can benefit the supply chain of physical products and service deliveries. According to Waller and Fawcett (2013) and Wang et al. (2016), BDA can provide value for the improvement of supply chains by maximizing productivity, improving operations and the decision-making process, reducing process variability and risks, support the demand and production planning, and choosing qualified suppliers.

However, it is not a trivial task to achieve those benefits. Many scholars have identified multiple challenges hindering the implementation of big data (Sanders, 2016). They also proposed frameworks to overcome these challenges (Dutta and Bose, 2015). For example, Chen and Zhang (2014) identified principles for designing big data systems and challenges of handling big data.

Although there is a consent among researchers on the positive impact of big data in SCM and the proposed approaches towards implementing big data appear reasonable and well-considered, many scholars did not back up their research through practical application, resulting in a lack of empirical foundation of their claims. The literature can provide the first evidence on the effectiveness of BDA in SCM. Then, a systematic literature review (SLR) was conducted to gather data about the BDA application in SCM.

This paper aims to review the academic literature on empirical studies concerning BDA implementation in SCM and to compare the results to theory. The comparison focuses on the challenges and benefits of each implementation. Therefore, the following research questions are addressed in this paper: What are the benefits and challenges of BDA implementation in SCM in conceptual and empirical studies?

This paper is organized as follows. Initially, the state-of-art of BDA application in SCM is present. Next section, "Research Design" gives insight into the SLR process, including search keywords, database, and inclusion criteria. The section "Big data analytics application in Supply Chain Management" contains the description of benefits and challenges related to BDA implementation in SCM, in theoretical and empirical works, resulted from the SLR. Additionally, similarities and differences between the theoretical and empirical findings, and future research suggestions are discussed in the section "Discussion and Future Research". Finally, the section "Final Remarks" offers the final considerations and outcomes of this work.

## Theoretical Background

The development of faster, smaller and cheaper computing devices has made digital communication ubiquitous in business (Addo-Tenkorang and Helo, 2016; Wamba and Akter, 2015). Such computational environment enabled the creation of new digital services and products for consumers, such as social media, cloud storages, Internet of Things (IoT) devices, online retailers, streaming services, and mobile apps (Brinch, 2016; Wamba and Akter, 2015). In the business environment, the digitalization is increasingly replacing the manual processes, and entirely new business models, which would not have been possible just a few years ago, are arising (Bharadwaj et al., 2013). An increasing amount of data have been generated and stored as a consequence of ongoing progress in information technology and digitalization. Diverse devices like sensors, smartphones, and computer systems increase the quantity of gathered data (Addo-Tenkorang and Helo, 2016).

Volume, velocity, variety, veracity, and value are known as the 5 Vs characterize big data. Volume refers to the massive amount of available data. Velocity refers to the speed of these data are generated, processed and analyzed. Variety is related to the different data types (e.g., structured and unstructured) which are collected from different sources. Veracity describes the reliability of data and in which extend data can be trusted. Finally, value can add value to the business (Demchenko et al., 2013; Wamba et al., 2015; Gandomi and Haider, 2015).

Is it widely recognized that big data contains valuable information that benefits a variety of domains and application areas (Yin and Kaynak, 2015). The benefit will come true when applies BDA in a large amount of data to derive actionable knowledgeBDA. The use of BDA can provide opportunities aiming to support faster and more data-driven decisions, leverage advanced analytics to measure trends more precisely, and optimize business processes (Hazen et al., 2016; Addo-Tenkorang and Helo, 2016; Brinch, 2016).

The SCM is one of the domains affected by the increase of digital infrastructure. A variety of data sources can provide more information about the supply chain operations, the market situation, and the financial activities of the company (Brinch, 2016; Jain et al., 2017). According to Hazen et al. (2014), "supply chain managers are increasingly reliant upon data to identify trends in costs and performance, and support process control and inventory monitoring".

Examples for applications of BDA in SCM include: the use of predictive maintenance to avoid unforeseen breakdowns (Jain et al., 2017); the tracking of delivery vehicles to identify and mitigate potential problems earlier (Mani et al., 2017); the prediction of price changes on products to adjust order size and order time (Chipizda et al., 2016); the performance measurement of current suppliers for sourcing decisions in the future (Wang et al., 2016); and the analysis of product usage from IoT devices at the customers location for marketing and product development purposes (Zaslavsky et al., 2013).

The findings from a Delphi study conducted by Brinch et al. (2017) "reveal that big data is more about data collection than data management and utilization, and that service, logistics, and planning processes are more applicable processes for adapting big data applications than manufacturing, sourcing, and return processes". Therefore, BDA application in SCM can help to increase efficiency and effectiveness in many areas of the supply chain to create a competitive advantage (Wamba and Akter, 2015).

**Research Design**

The research design was an SLR application in two stages. First, the authors have looked up the literature reviews to identify the theoretical and technological aspects of BDA for SCM. Instead of developing an SLR, the authors decided to take advantage of a tertiary study conducted by Mello et al. (2017) on six literature review papers. We updated Mello et al. (2017) adding Tiwari et al. (2018). So, seven theoretical papers, which conducted a literature review of big data applications in SCM, were analyzed.

The purpose of the second stage was to identify academic papers with empirical studies of BDA implementation in SCM. The SLR followed the steps proposed by Thomé et al. (2016): planning and formulating the problem; searching the literature; data gathering; data analysis and synthesis; interpretation; and presenting results. The outcome is empirical papers related to BDA for SCM.

The search, realized in January 2018, selected papers which have the combination of the terms big data, supply chain or logistics, and one of the terms related to empirical studies (i.e., implementation, implementing, application or empirical) in the field of title

or abstract. Only journals and conference papers written in English were selected. The Scopus database was selected because Mishra et al. (2016) argue such database includes over 20,000 peer-reviewed journals and index relevant papers for the area of supply chain and big data. In total, the initial search resulted in 168 papers.

The authors read and analyzed all title and abstract of selected papers aiming to find those whose potentially describe the experiences of implementing BDA in SCM. The inclusion criteria were papers which included empirical studies that are thoroughly documented and contain descriptions of the challenges and benefits for implementing BDA in SCM. Six articles filled the requisites for further investigation. Those papers contain empirical findings from three case studies, one survey, and two Delphi studies.

## Big Data Analytics Application in Supply Chain Management

This section first provides an overview of the BDA benefits and challenges in SCM gathered from the theoretical and empirical studies found in the two stages of the SLR. Afterward, the similarities and differences between the theoretical and empirical findings are discussed in the Section "Discussion and future research".

### *Theoretical Studies*

All authors of seven identified theoretical papers agree that the BDA implementation in SCM will lead to an improved capability to make decisions. The BDA application can enhance speed and flexibility of making decisions. The real-time data streams can improve the decision response time. Finally, these decisions can be more strategic data-oriented (Addo-Tenkorang and Helo, 2016; Tiwari et al., 2018; Brinch, 2016). The improved decision-making capability will realize benefits for SCM on all levels. The different decision levels, i.e., strategic, tactical and operational, made on a day-to-day basis are affected, leading to a more effective and efficient SCM (Mishra et al., 2016; Wamba and Akter, 2015).

On a strategic level, BDA can help designing a new supply chain by selecting locations for warehouses through the analysis of transportation and customer data as well as the improvement of strategic sourcing by assessing and selecting the best potential suppliers (Tiwari et al., 2018; Wang et al., 2016). Data from various sources, like loyalty card programs, social media activity, sales data or IoT data generated by their products, can improve the understanding of customers' demands providing insightful information to development of new products and services or create personalized offers (Mishra et al., 2016; Brinch, 2016; Wamba and Akter, 2015). On a tactical level, the newly available data also enables the definition of new and more detailed performance metrics that managers can use to closely monitor both the supplier and employee performance to take appropriate actions (Addo-Tenkorang and Helo, 2016; Mishra et al., 2016; Tiwari et al. 2018). Monitoring the supply chain more closely also helps to identify potential fraud or malfeasance and bottlenecks (Wamba and Akter, 2015). The demand planning can also be enhanced using improved forecasting, and the real-time traffic data can incorporate into delivery planning for reducing carbon emission (Tiwari et al., 2018).

Overall, BDA in SCM leads to an enhancement of productivity (Mishra et al., 2016), a more agile supply chain that flexible and responsive to changes and problems (Rozados and Tjahjono, 2014; Wand et al., 2016), a higher level of end-to-end visibility (Mishra et al., 2016; Tiwari et al., 2018), and improved supply chain stakeholder relationships (Mishra et al., 2016; Wamba and Akter, 2015).

There are challenges to achieve those benefits and successfully implement BDA in SCM. Due to the massive changes in SCM required by the BDA implementation, the

active involvement of all supply chain stakeholders to change the organizational culture and the strategy to the new circumstances is imperative (Rozados and Tjahjono, 2014; Wang et al., 2016). Furthermore, the employees who have adequate analytical capabilities to conduct the BDA are essential and hard to find (Tiwari et al., 2018). Before the BDA implementation, integrating the organizational processes, information systems and possible data silos within single departments are required to reduce redundancies in development and maintenance efforts (Wamba and Akter, 2015; Rozados and Tjahjono, 2014). From a technological point of view, it is also a challenge to implement BDA because it requires advanced hardware and software, and complex architecture of algorithms (Tiwari et al., 2018). Finally, the big challenge for a successful implementation is the validation of theoretical frameworks that decrease the risk of failures (Addo-Tenkorang and Helo, 2016; Rozados and Tjahjono, 2014; Brinch, 2016).

***Empirical Studies***

After applying the SLR, we selected six papers which posed empirical evidence related to BDA application in SCM. Mani et al. (2017) explain the implementation of a delivery vehicle tracking system in an Indian co-operative milk-processing organization. Chipidza et al. (2016) describe the case of a grocery wholesaler pursuing the BDA implementation for price prediction within their supply chain. Dutta and Bose (2015) conducted a case study in a company of the cement manufacturing industry. Furthermore, Schoenherr and Speier-Pero (2015) present the results of a survey with practitioners that already worked on implementing big data in SCM. Kache and Seuring (2017) carried out a Delphi study among 20 consultants with a background either in technology or management. Finally, Jain et al. (2017) also employed Delphi technique aiming to identify issues of SCM which can be resolved using BDA.

BDA can be used during the design phase to build a supply chain that is more collaborative and integrated with its partners. At runtime, the available data can be analyzed to improve the understanding of the customer requirements to improve the products or to introduce new, innovative solutions (Kache and Seuring, 2017). Demand data can be used to identify underdeveloped markets or improve customer segmentation to provide personalized offers and better target individual marketing efforts at customers, leading to improved sales at lower marketing prices (Dutta and Bose, 2015; Jain et al., 2017; Kache and Seuring, 2017).

Overall, the availability of more information leads to more data-oriented, informed decision-making (Dutta and Bose, 2015; Schoenherr and Speier-Pero, 2015; Jain et al., 2017). The availability of large amounts of data from multiple sources allows for the definition of new metrics, enabling improved monitoring of the entire supply chain (Mani et al., 2017; Dutta and Bose, 2015). Due to a more granular and timely overview of deliveries, production planning, and demand, BDA helps to optimize inventory levels and improving the availability of goods while reducing costs at the same time (Schoenherr and Speier-Pero, 2015; Jain et al., 2017; Kache and Seuring, 2017). Furthermore, the availability of real-time operational data allows for the timely detection and prevention of unwanted events that could lead to expensive repercussions (Mani et al., 2017; Kache and Seuring, 2017).

Generally, BDA enables the improvement of supply chains' operational efficiency (Schoenherr and Speier-Pero, 2015; Kache and Seuring, 2017) and a reduction of costs through a speedup of unload times, a reduction of penalties, damages or demurrages (Dutta and Bose, 2015), the ability to perform real-time route rescheduling, and an improved understanding of how the operational practice impacts the overall costs

(Kache and Seuring, 2017) and customer satisfaction (Dutta and Bose, 2015). Predicting price developments enables resellers to pick the right time to order to increase their profit margin (Chipidza et al., 2016). BDA in SCM enhances the visibility and responsiveness to change and problems (Schoenherr and Speier-Pero, 2015; Kache and Seuring, 2017). Besides, BDA enables predictive maintenance which can prevent unnecessary and unplanned maintenance (Jain et al., 2017).

There is a bundle of challenges to be overcome to achieve these benefits. Introducing a BDA in a company requires employees to adapt and use the new tool in their daily working practice. They need to develop a mindset that focuses data and facts. These changes within the culture must be managed and enforced. Using BDA in a supply chain across functional departments or firms also requires a high level of collaboration and a change in organizational structure (Dutta and Bose, 2015; Schoenherr and Speier-Pero, 2015; Kache and Seuring, 2017). Introducing the new software solutions is also a high technical barrier. There might be a lack of a robust IT infrastructure, or the existing one is unstandardized and fragmented and needs to be integrated to make all relevant information available throughout all IT systems. Therefore, massive and expensive changes in the infrastructure might be necessary to prepare it for the introduction of the analytical system.

The analytical system availability requires expensive investments (Schoenherr and Speier-Pero, 2015; Kache and Seuring, 2017). Also, an appropriate technological solution might be unavailable at the market (Chipidza et al., 2016; Schoenherr and Speier-Pero, 2015). Using the system requires a high level of analytical and technological expertise which imposes a challenge on balance human and analytics management style (Schoenherr and Speier-Pero, 2015; Kache and Seuring, 2017). Other factors that could inhibit the decision for an introduction of BDA in SCM are time constraints, the lack of a clear business case or the impression of stakeholders and managers that introducing and managing analytics in the company is overwhelming and too complicated (Schoenherr and Speier-Pero, 2015; Kache and Seuring, 2017). After the introduction of the analytical system, governance and compliance rules need to be created and enforced by the company. The complexity of the newly available information requires careful information management. The information security becomes a greater issue due to the deeper insights into the company that could also cause harm if it became available to unauthorized parties. (Kache and Seuring, 2017). Finally, even if BDA provides great insight into every detail of the supply chain, managers need to be able to balance a data-driven and a human-oriented managerial style (Kache and Seuring, 2017).

**Discussion and Future Research**

Comparing the potential benefits foreseen by the authors of theoretical papers to those by authors of the empirical papers, we come up with some clear matches. Both sets of authors argue the companies which implement BDA in their SCM will obtain an improved data-driven decision-making capability. Empirical data also support the hypothesis the BDA provides a better understanding of customers, enabling the new products and services development adjusted to the customer's demands. Marketing profits from the ability to create market niches and precisely appeal to individual customers through targeted advertisement and offers. The empirical findings support the hypothesis the newly available data might be used to create new metrics to support the SCM to improve the operational efficiency. The availability of real-time information can indeed reduce the time to react to ongoing events to prevent undesired consequences such as delays because of traffic. Conceptual and empirical works

highlight the benefits such as increased visibility, cost reduction, and improved demand planning.

The predictive maintenance to avoid breakdowns and reduce maintenance costs is an important benefit only mentioned by the authors of empirical. Another benefit, according to only to the authors of theoretical papers, is the possibility to identify the optimal warehouses' locations and suppliers when designing a supply chain. The empirical data, however, show the BDA implementation across firms during the design phase of supply chain improves the collaboration and integration of the involved companies.

Regarding the main challenges to implement BDA in SCM, the authors of the empirical papers have been facing the same threats by the authors of the theoretical papers. The implementation requires extensive changes in culture, organizational structure, governance, and processes. Further, integrating and extending the existing IT infrastructure require high investments even before the BDA itself is introduced. Finally, it is hard to find the required skilled employees with a technological and analytical background.

The authors of empirical papers, especially Kache and Seuring (2017), highlight more challenges than the ones of theoretical papers. It seems the reality of implementing BDA is very complicated. Hence, further research on the barriers and challenges of implementing BDA appears to be a promising and beneficial endeavor (Rozados and Tjahjono, 2014). The future studies could enrich the findings of this paper by conducting case studies among companies which use BDA in SCM to verify the benefits and challenges to enrich the studied field. As the BDA application in SCM is a recent topic and the interest in it has been increasing since 2013 (Mishra et al., 2016), we recommend repeating this investigation one or two years later aiming to include new theoretical and empirical works to increase the knowledge of the applications of BDA in SCM.

**Final Remarks**

Although many papers are focusing on the topics BDA and SCM and a high interest in those issues, the empirical studies remain sparse. From the 168 identified papers during the SLR process, only six papers contain detailed information about the benefits and challenges of BDA in SCM from real-world experiences. This paper compared the findings of theoretical and empirical papers. It is necessary to consider studies in those topics conducted together with practitioners, for example, case studies and Delphi studies, to fill the gap of information.

Only Dutta and Bose (2015) provided both a comprehensive framework on how to implement big data in SCM and tested it in a real-world case. So, it is necessary more implementations to generalize the findings. Dutta and Bose (2015) emphasize the need for a clear roadmap that combines both data management and change management to enable an organization to not only use the required software but also to make use of big data in their day-to-day business actively.

This paper identified several benefits for SCM provided through the implementation of BDA as increased visibility, reduce costs, and demand planning. Likewise, there are challenges as information technology infrastructure and acceptance by employees. Therefore, this paper provides insights from empirical and theoretical studies regarding the benefits and challenges using BDA in SCM.

## Acknowledgments

The authors acknowledge the Brazilian National Council for Scientific and Technological Development (CNPq) - Science without Borders Program for the support with a research grant.

## References

Addo-Tenkorang, R. and Helo, P. T. (2016), 'Big data applications in operations/ supply-chain management: A literature review', *Computers & Industrial Engineering*. 101, pp.528-543.

Bharadwaj, A., El Sawy, O., Pavlou, P., & Venkatraman, N. (2013). Digital business strategy: toward a next generation of insights.

Brinch, M. (2016), 'Big data and supply chain management: A content-based literature review', in *The 23rd EurOMA conference*, pp. 1–13.

Brinch, M., Stentoft, J., & Jensen, J. K. (2017), 'Big data and its applications in supply chain management: Findings from a Delphi study' In *Proceedings of the 50th Hawaii International Conference on System Sciences*.

Chen, C. P. and Zhang, C.-Y. (2014), "Data-intensive applications, challenges, techniques and technologies. A survey on Big Data", *Information Sciences*, Vol. 275, pp. 314–347.

Chipidza, W., George, J. and Koch, H. (2016), "Chartering Predictive Analytics. A Case Study".

Demchenko, Y., Grosso, P., De Laat, C., & Membrey, P. (2013). 'Addressing big data issues in scientific data infrastructure', In *International Conference on Collaboration Technologies and Systems*, IEEE, pp. 48-55.

Dutta, D. and Bose, I. (2015), 'Managing a big data project. The case of Ramco cements limited', *International Journal of Production Economics*, Vol. 165, pp. 293–306.

Gandomi, A., & Haider, M. (2015). 'Beyond the hype: Big data concepts, methods, and analytics', *International Journal of Information Management*, 35(2), 137-144.

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', *International Journal of Production Economics*, 154, pp.72-80.

Hazen, B. T., Skipper, J. B., Boone, C. A., & Hill, R. R. (2016). 'Back in business: Operations research in support of big data analytics for operations and supply chain management', *Annals of Operations Research*, pp. 1-11.

Jain, A. D. S., Mehta, I., Mitra, J., and Agrawal, S. (2017), 'Application of Big Data in Supply Chain Management', *Materials Today: Proceedings*, *4*(2), pp.1106-1115.

Kache, F. and Seuring, S. (2017), 'Challenges and opportunities of digital information at the intersection of Big Data Analytics and supply chain management', *International Journal of Operations & Production Management*, Vol. 37 No. 1, pp. 10–36.

Mani, V., Delgado, C., Hazen, B.T. and Patel, P. (2017), 'Mitigating Supply Chain Risk via Sustainability Using Big Data Analytics. Evidence from the Manufacturing Supply Chain', *Sustainability*, Vol. 9 No. 4.

Mishra, D., Gunasekaran, A., Papadopoulos, T. and Childe, S.J. (2016), 'Big Data and supply chain management. A review and bibliometric analysis', *Annals of Operations Research*, pp. 1–24.

Mishra, D., Luo, Z., Jiang, S., Papadopoulos, T., & Dubey, R. (2017). 'A bibliographic study on big data: concepts, trends and challenges', *Business Process Management Journal*, 23(3), 555-573.

Mello, R., Scavarda, L. F., Hellingrath, B., Martins, R. (2017) 'Big data and data analytics for Supply Chain Management: a tertiary study', *NOFOMA 2017*, pp.490–503.

Rozados, I.V. and Tjahjono,B. (2014) 'Big data analytics in supply chain management: trends and related research', *6th International Conference on Operations and Supply Chain Management*. Bali.

Russom, P. (2011). 'Big data analytics'. *TDWI best practices report*.

Sanders, N. R. (2014), *Big data driven supply chain management: A framework for implementing analytics and turning information into intelligence*, Pearson Education.

Sanders, N. R. (2016), 'How to use big data to drive your supply chain', *California Management Review*, 58(3), 26-48.

Schoenherr, T. and Speier-Pero, C. (2015), 'Data science, predictive analytics, and big data in supply chain management. Current state and future potential', *Journal of Business Logistics*, Vol. 36 No. 1, pp. 120–132.

Thomé, A. M. T., Scavarda, L. F. and Scavarda, A. J. (2016) 'Conducting systematic literature review in operations management', *Production Planning & Control*, 27 (5), pp.408–420.

Tiwari, S., Wee, H.M., Daryanto, Y. (2018) 'Big data analytics in supply chain management between 2010 and 2016: Insights to industries'. *Computers and Industrial Engineering*, Vol 115, pp. 319–330.

Waller, M.A., Fawcett, S. E. (2013) 'Data Science, Predictive Analytics, and Big Data: A revolution that will transform supply chain design and management', *Journal of Business Logistics*, Vol.34, pp.77-84.

Wamba, S.F. and Akter, S. (2015) 'Big Data Analytics for Supply Chain Management:A Literature Review and Research Agenda', *Workshop on Enterprise and Organizational Modeling and Simulation*. Springer International Publishing, pp.61–72.

Wamba, S.F., Akter, S., Edwards, A., Chopin, G. and Gnanzou, D. (2015), "How 'big data' can make big impact: Findings from a systematic review and a longitudinal case study", *International Journal of Production Economics*, Vol. 165, pp. 234–246.

Wang, G., Gunasekaran, A., Ngai, E.W.T. and Papadopoulos, T. (2016), "Big data analytics in logistics and supply chain management. Certain investigations for research and applications", *International Journal of Production Economics*, Vol. 176, pp. 98–110.

Yin, S., & Kaynak, O. (2015). 'Big data for modern industry: challenges and trends [point of view]'. *Proceedings of the IEEE*, 103(2), pp, 143-146.

Zaslavsky, A., Perera, C., & Georgakopoulos, D. (2013). 'Sensing as a service and big data'. arXiv preprint arXiv:1301.0159.