Big Data-Enabled Supply Chain Management: A systematic review

Roy Meriton and Gary Graham

1. Introduction

Big data has the potential to revolutionize the art of supply chain management (Mehmood et al., 2016). Reporting on the economic impact of big data (Manyika et al., 2011), the McKinsey Global Institute has identified value levers along the whole manufacturing value chain with the ability to lead to high productivity levels. Application of big data and advanced analytics to R&D and product development has been estimated to reduce costs by 20-50 per cent, and a big data-enabled supply chain optimization was estimated as yielding a 2-3 percentage point profit margin improvement (Baily et al., 2013). However, in spite of its widely reported strategic impacts, there is a paucity of empirical research exploring the influence of big data in operations and supply chain management (SCM).

Fosso Wamba et al. (2015, p235) define big data as:

… a holistic approach to manage, process and analyse the 5Vs (volume, velocity, variety, veracity, and value) in order to create actionable insights for sustained value delivery, measuring performance and establishing competitive advantages.

Big data analytics (BDA), on the other hand, is the process of using analysis algorithms running on powerful supporting platforms to uncover potentials concealed in big data, such as hidden patterns or unknown correlations (Hu et al., 2014). Hence, as Russom (2011) observes, BDA is really about two things – big data and analytics – plus how the two have teamed up to create one of the most profound trends in business intelligence (BI) today. According to Fosso Wamba and Akter (2015), BDA is expected to take supply chain (SC) transformation to a level of transformation never before achieved. This could be because BDA represents a critical source of meaningful information that may help SC stakeholders gain improved insights that can be used for competitive advantage. For example, Fawcett and Waller (2014) expose how Amazon is making use of the predictive capability of BDA to predict and ship what a customer wants before the customer places their order. This has been termed anticipatory shipping (Lee, 2016). In turn, this can improve the customer experience and thus a source of competitive advantage. However, whilst supply chain professionals are devising new and innovative ways to extract value from BDA-enabled activities, there is a dearth of empirical academic work published in the last five years, particularly in the business and
management subject area. This may be because the area is still an emerging field early in its developmental stage.

The aim of this work is therefore to examine the academic literature with a view of mapping the common terrain where BDA and SCM intersect in a bid to advance knowledge on the emerging dynamics in this space. We are thus guided by the following research questions:

- What is the potential of BDA in the SCM to improve performance?
- Are there any barriers to adopting BDA in SC, and if so what are they?

We conducted a detailed systematic survey of the literature in the traditional manner using keywords, complemented by visualisations in VOS MAX and cluster analysis of the most common terms from the articles extracted.

2. Methodology

In this study we build on the previous review conducted by Fosso Wamba and Akter (2015) which was based on a hybrid approach. This approach consisted of a search using the following keywords: ‘Big data’ AND ‘supply chain’ within the Scopus database. Scopus, officially named SciVerse Scopus, was introduced by Elsevier in November 2004 to the information market (Aghaei Chadegani et al., 2013). It is the largest database existing on the market for multidisciplinary scientific literatures (Bornmann et al., 2009) and has often been used as a valuable information repository for supply chain scholars (e.g. Fahimnia et al., 2015). Scopus covers 27 million abstracts, 230 million references and 200 million web pages (Bar-Ilan, 2008). It contains 20,500 peer-reviewed journals from 5,000 publishers, together with 1200 Open Access Journals, over 600 Trade Publications, 500 Conference Proceedings and 360 book series from all areas of science (Aghaei Chadegani et al., 2013).

The search was conducted on Monday 19 September 2016, and our search returned 139 titles registered in the Scopus database. These included 59 journal articles of which 12 were in press. Conference papers were in the majority with 75 entries, the balance being made up of reviews, one book chapter and a short survey. We subjected the articles from the result of the Scopus search to a cluster analysis to help provide a holistic picture of the most common words/themes and the relationship between them. The visualizations of these are illustrated in Figure 1.
3. Result

This section presents and discusses the findings of big-data-related SC publications identified in Scopus.

In Table 1, the distribution of publications by the year of publication is presented. The trend suggests an increasing interest in the subject area with 2015 accounting for almost 50% of all the publications (46%). This upward intake in trend is also apparent with the articles, but it remains to be seen whether 2016 will be as successful a year for the subject area as 2015.

**Table 1: Distribution of publications by year**

<table>
<thead>
<tr>
<th>Year</th>
<th>Number of Publication</th>
<th>Percentage</th>
<th>Number of Articles (incl. in press)</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>2012</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>2013</td>
<td>9</td>
<td>7</td>
<td>4</td>
<td>7</td>
</tr>
<tr>
<td>2014</td>
<td>28</td>
<td>20</td>
<td>9</td>
<td>15</td>
</tr>
<tr>
<td>2015</td>
<td>64</td>
<td>46</td>
<td>28</td>
<td>47</td>
</tr>
<tr>
<td>2016</td>
<td>36</td>
<td>26</td>
<td>17</td>
<td>29</td>
</tr>
<tr>
<td>Total</td>
<td>139</td>
<td>100</td>
<td>59</td>
<td>100</td>
</tr>
</tbody>
</table>

The distribution of publications by subject area is presented in Table 2. The top 4 subject areas accounted for 77% of all the publications and these include: (1) Computer Science (29%); (2) Engineering (24%); (3) Business, Management and Accounting (13%) and; (4) Decision Sciences (11%). It is interesting to note that the same four subject areas occupied the top spots in terms of most number of publications in Fosso Wamba and Akter's (2015) review, however in a slightly different order with 'Business, Management and Accounting' holding the top spot (25.8%). These findings indicate that the dynamics between supply chain and big data is being discussed from a more technological perspective rather than in a business and management sense, which is the traditional home to supply chain management. This may be attributed to the nascent nature of BDA which still commands a high cost for adoption.

**Table 2: Distribution of publications by subject area**

<table>
<thead>
<tr>
<th>Subject Area</th>
<th>Number of publications</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Computer Science</td>
<td>85</td>
<td>29</td>
</tr>
<tr>
<td>Engineering</td>
<td>69</td>
<td>24</td>
</tr>
<tr>
<td>Business, Management and Accounting</td>
<td>38</td>
<td>13</td>
</tr>
<tr>
<td>Decision Sciences</td>
<td>31</td>
<td>11</td>
</tr>
</tbody>
</table>
Table 3 shows the classification of publications by country. So far, most published articles on big data in the supply chain context come from China (24.2%), with the US following closely behind (22%). These figures are contrasted with the findings in Fosso Wamba and Akter (2015), who reported that the US authors held a significantly larger share of the publication spoils (63.2%) compared to the Chinese authors (21.1%). It seems that in just over one year we have seen a ‘change of the guard,’ with the rest of the world, spearheaded by China, also wanting to flex their muscles in this space.

Fosso Wamba and Akter (2015, p4) noted that:

\[\text{\ldots the impact of BDA in emerging economies and less developed countries should be part of future research directions.}\]

Indeed, the advice appears to have been heeded, with authors from countries such as Brazil, India, Russia, South Africa, and Pakistan beginning to make their voices heard, albeit respectively contributing to 1% of the total amount of publications. Nevertheless the lack of studies with a focus on small island states (SIDS) where the economy is predominantly dependent on tourism and fishery, such as Seychelles, suggests the untapped potential of this emerging literature. It is, however, not surprising that the most prolific authors researching in this space originate from the US and China respectively. Both S.E. Fawcett and M.A. Waller have 4 articles each on BDA-related SC topics to their names, followed closely by Li, P., Liu, Y.Q., and Luo, H., each recognised for 3 publications.
Table 3: Distribution of publications by country

<table>
<thead>
<tr>
<th>Country/Territory</th>
<th>Number of publications</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>China</td>
<td>41</td>
<td>24</td>
</tr>
<tr>
<td>United States</td>
<td>38</td>
<td>22</td>
</tr>
<tr>
<td>Germany</td>
<td>15</td>
<td>9</td>
</tr>
<tr>
<td>United Kingdom</td>
<td>9</td>
<td>5</td>
</tr>
<tr>
<td>India</td>
<td>7</td>
<td>4</td>
</tr>
<tr>
<td>Taiwan</td>
<td>5</td>
<td>3</td>
</tr>
<tr>
<td>Netherlands</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>Spain</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>Finland</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>France</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>Ireland</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>Singapore</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>South Korea</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>Australia</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>Brazil</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>Denmark</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>Hong Kong</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>Italy</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>Switzerland</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>Belgium</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Canada</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Czech Republic</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Greece</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Japan</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Malaysia</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Norway</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Pakistan</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Russian Federation</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>South Africa</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Sweden</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Undefined</td>
<td>12</td>
<td>7</td>
</tr>
</tbody>
</table>

*Total 173 100

Note: *Some articles are counted more than once because their authors come from different countries

4. Findings

We elected to retain the entries classed as ‘journal articles’ in the database for a detailed content analysis. Following further screening of the abstracts, a total of 35 articles were found to meet our criteria for inclusion for detailed analysis. The main reason articles
were excluded from the final list related to whether or not the full text was available online. From the sample of 59 journal articles, 17 were dropped because the full text document could not be located or downloaded; 2 were excluded as they were not scholarly articles; and a further 5 were deemed not relevant to the study. The analysis of the 35 articles captured attributes of the studies such as methodological approaches and the industries. The graphics in Figure 1 provide a snapshot of the most significant relationships. In the sections that follow we draw on the analyses in Figure 1 to discuss the headlines of our findings, paying particular attention to emerging trends; the most common of these include method, opportunities and barriers.

**Figure 1: Visuals of dominant themes and relationships**

4.1 Dominant methodological approaches

The methodological approach was a mixed bag of quantitative methods and literature reviews often leading to a conceptual and/or theoretical framework. In the main, the studies analysed were finely poised with 18 being mostly quantitative and 17 non-quantitative in nature. There seems to be an increasing level of novelty and complexity in the methods employed. Perhaps this is a reflection on the nature of the task. For instance, the quantitative studies were often preceded by mathematical modelling followed by experiments or case studies to test the model derived. For example, Bock and Isik (2015) developed a mathematical model to quantify *intricacy* in purchasing order
sizing processes. They then followed this up by using ‘structured’ big data related to costs and the structure of recent deviations from desired target values, to conduct scenario analyses in order to determine the reliability of the model. Only 4 studies relied on primary quantitative data collected by means of survey questionnaires, these include: Chen et al. (2015) who surveyed 161 U.S.-based companies; van der Spoel et al. (2015) who collected data from 230 truckers; and Schoenherr and Speier-Pero (2015) whose study involved 531 SCM professionals. The other study by Papadopoulos et al., (2015) made use of both survey data as well as data from social media, such as Twitter and Facebook. However, most of the studies actually made use of secondary ‘big data’.

Big data employed was of both the structured and unstructured nature. In Li et al. (2015) the authors analysed structured data obtained from an omni-channel marketing platform to test a new enterprise networks integration architecture. On the other hand, Papadopoulos et al. (2015) analysed data in the form of tweets and Facebook posts, to help develop a framework for supply chain resilience. Other means of gathering big data included RFID-enabled intelligence from a shop floor environment (Zhong et al., 2016). The non-quantitative method (as we have termed it) was dominated by conceptual and/or theoretical framework development often involving a systemic survey of the literature. It is interesting to note that only one article in this category actually employed a qualitative approach of the conventional type. In this study Kumar et al. (2016) adopted a case study approach at the product level to gauge the impact big data was having on the traditional supply chain.

4.2 Opportunities presented by BDA to SCM

Our findings suggest that the potential of BDA as a valuable resource to the organization has not gone unnoticed. However, there are two sides to this ‘opportunities coin’; this is because, there is much by way of commentaries, theoretical framing, simulations and experiments and still little empirical evidence. Much of the conversation tends to be past oriented, in the sense that authors consider existing supply chain analytics and speculate on how the integration of BDA could potentially help to optimise supply chain activities. A good example is provided by Rehman et al. (2016). In this work the authors introduce the idea of early big data reduction (BDR) and go on to make a case of how it can be leveraged in supply chains to improve efficiency. Antagonism about the lack of empirical work aside, there is widespread sentiment that BDA is a necessary evil that, increasingly, firms ignore at their own peril. In this context, the value potential of BDA unearthed by the findings can be themed around (1) capabilities and, (2) performance.

4.2.1 Capabilities

Drawing on the findings of the empirical studies, scholars generally recognised that BDA could be leveraged in different parts of the supply chain in order to create value. In other words, there is widespread acceptance that BDA is valuable when it is used to create
distinct capabilities. In the main, the findings reveal that organizations are making use of the predictive proclivities of BDA to strengthen their decision-making capabilities (Schoenherr & Speier-Pero, 2015) in a number of key supply chain activities. Some of these big-data enabled capabilities include market sensing (Chae, 2015; Lee, 2016; Li et al., 2015; Liu & Wang, 2016), planning and forecasting in different areas such as in logistics (Liu & Wang, 2016; Zhong et al., 2015) and demand and sales (Schoenherr & Speier-Pero, 2015), risk management (Papadopoulos et al., 2015; Wu et al., 2015; Zhao et al., 2015; Zou et al., 2016) and innovation (Tan et al., 2015).

4.2.2 Performance
Distinct capabilities often bestow performance advantages on the incumbent firms. The BDA-enabled performance advantages revealed by the findings include cost efficiencies (Bock & Isik, 2015; Hofmann, 2015; Li & Wang, 2015; Li et al., 2015; Liu et al., 2016; Schoenherr & Speier-Pero, 2015), enhanced customer services (Lee, 2016; Liu & Wang, 2016), agility, in terms of speed, flexibility and responsiveness (Giannakis & Louis, 2016; Kumar et al., 2016) and business growth (Chen et al., 2015).

4.3 Barriers to adopting BDA in SCM

Whilst the value of BDA to firms was recognised, this was also greeted by an air of cautious optimism by many authors. The conversations around barriers in implementing BDA in SCM mostly took place in the non-empirical articles. Cost associated with currently available solutions (Schoenherr & Speier-Pero, 2015) was one of the most cited drawbacks, for example, Rehman et al. (2016) observed that cloud service utilization costs increase because of big data analytics and value creation activities. On a related note, Hazen et al. (2014) pondered on data quality noting that the costs of poor data quality have been estimated to be as high as 8% to 12% of revenues for a typical organization and may generate up to 40% to 60% of a service organization's expenses. Issues around organizational readiness have also been highlighted. For instance, Schoenherr and Speier-Pero (2015) found employees being inexperienced, time constraints, lack of integration with current systems, change management issues, lack of appropriate predictive analytics solutions for SCM, as well as the perception of SCM predictive analytics being overwhelming and difficult to manage amongst the primary barriers for adopting BDA in SCM. It is therefore not surprising that Sonka (2014) concluded that successful application of big data and associated analytical techniques in the agriculture supply chain would also depend on organizational and managerial factors.

5. Discussion and conclusion

In this paper we sought to understand the potential of BDA in the SC to improve performance, as well as the barriers, if any, that constrain its adoption in SC. We
conducted a detailed literature survey using ‘big data’ and ‘supply chain’ as key words. The search returned 139 publications. Applying VOS MAX to the search results we complemented the initial search with visualisations and cluster analysis that helped us understand the dominant themes and the nature of the relationships between them. Thus we provided a robust approach to analyse the existing literature.

Our analysis suggests an increasing trend in the number of publications, year on year, with authors from more countries getting involved. Particularly, the last five years have seen a shift from the US towards the east in terms of the number of articles published, with an increasing participation of authors from emerging economies such as Brazil, Russia and India, with China leading the way. Perhaps this trend could be attributed to faster uptake of BDA in SCM particularly in China, which in turn could reflect the relative costs of such technological adoption in China compared to the US and the rest of the world for that matter. Furthermore, the findings reveal that 'business and management' area is lacking behind computer science and engineering in terms of representativeness. This could be explained by the fact that most of the empirical work is of a quantitative nature often involving complex modelling using advanced algorithms which make use of secondary data. Perhaps scholars from these specialisms, that is, engineering and computer science, are better equipped with the relevant skills for these approaches. The dominant methodologies in the business and management discipline generally revolve around primary data collection through surveys or case studies, both of which require greater effort to collect data while access to companies, especially in the western economies, is becoming much harder to secure.

The findings suggest that adoption of BDA in the SC has the potential to impact on firm performance. However, we found that the value of BDA to firms is not inherent in the BDA per se but rather by being embedded within supply chain activities. In other words, BDA is causally efficacious in enabling the emergence and evolution of unique capabilities. In particular, the findings reveal that firms are leveraging the predictive proclivities of BDA to strengthen their decision-making capability for various SC activities; here we wish to single out sensing and innovation capabilities. These capabilities are quintessential for competing in dynamic environments. For example, sensing capabilities feature as one of the cornerstones of Teece’s (2007) microfoundations of dynamic capabilities (DC). Furthermore, innovation is often cited as the linchpin of dynamic competition (Schumpeter, 2000; Zollo & Winter, 2002). According to Eisenhardt and Martin (2000, p1107), dynamic capabilities are:

… the firm’s processes that use resources – specifically the processes to integrate, reconfigure, gain and release resources – to match and even create market change.

The emergence of anticipatory shipping, based on BDA-enabled sensing capabilities from Amazon can be thus be seen in the light of creating market changes which therefore chimes with the idea of DC. Here we wish to press home the important point that DC is
a strategic enterprise. Hence, the findings serve to reiterate the strategic importance of a firm’s supply chain particularly as a means of sustainable competitive advantage (Opresnik & Taisch, 2015; Rehman et al., 2016) and firm growth (Chen et al., 2015). Therefore, the findings serve to position BDA in a firm’s strategic conversation particularly illuminating its role as an enabler of dynamic capabilities and sustainable competitive advantage.

The findings also reveal that the adoption of BDA improves supply chain agility, particularly noticeable in performance metrics of flexibility, responsiveness and speed (Giannakis & Louis, 2016). These operations performance measures are linked with supply chain resilience. In fact, a number of authors have accentuated the criticality of BDA to SC risk management with Wu et al. (2015) arguing that applying relevant analytical techniques to unstructured and structured data can help firm identify areas of risk that may impact on the sustainability of the supply chain. It is now widely accepted that SC resilience can be a dynamic capability. However, whilst DC tend to focus on a particular functional area, such as R&D for instance, resilience speaks to a whole system’s approach to adaptation and can be both proactive and reactive. Supply chain resilience is widely considered as an important component of supply chain management, however, it is surprising that only two articles make direct reference to it. Nevertheless, resilience presents firms with a different route to adaptation and the findings provide some useful insights as to the dynamics of BDA-enabled resilient supply chains.

5.1 A look into the future

The litmus test for adopting BDA in SCM rests on a number of permutations; accessibility being one of them. The findings reveal cost of the available options as one of the key barriers. This scenario tends to put particularly small firms at a disadvantage, in fact, it seems to be the case that the uptake of this technological innovation is a lot more celebrated in large and well established firms; Amazon, Ikea, Toyota, to name but a few. The cost of adoption may also be one of the reasons accounting for the surge of publications from China. It may be the case that the cost of adoption in China is relatively cheaper than in other parts of the world encouraging a higher level of adoption. This could potentially result in a wider caucus of firms making use of BDA in their operations and thus providing for a larger population of organizations for researchers to draw from. Whilst we agree with Fosso Wamba and Akter (2015) in that BDA has the potential to take supply chain (SC) transformation to a level never before achieved, barriers such as costs and the organizational readiness issues highlighted would need to be addressed a priori.

Furthermore, BDA is a relatively recent addition in the SCM domain and perhaps still exudes an air of cautious optimism. More needs to be done in terms of researching cost-effective solutions that are currently being used. Most of the empirical studies that we have analysed tend to be detached from real life settings. There is a real world out there
where these technologies are being put to use on a daily basis and companies are reaping the rewards. We are thus advocating for research of the longitudinal case study type, perhaps of an ethnographic approach, in order to further our understanding as to the process of capabilities emergence and evolution in the context of BDA. It is our conviction that by doing so scholars will help to disseminate best practices more widely which in turn could increase the rate of adoption and bring down the costs involved whilst also exposing the benefits to a wider audience.

We are also convinced of the untapped potential of big data technologies for small island developing states. We can think of a few areas where these could be leveraged to improve the competitiveness, sustainability and indeed resilience of supply chains. For example, the danger of wind shear is a major concern for the civil aviation authority in Seychelles and often flights have been diverted, especially during the trade winds, to neighbouring islands such as Mauritius. These instances carry a cost element, not just for the authority but also through the ripple effect on the tourism supply chain. Weather data can help to optimise route planning to avert instances of heavy wind shear loadings during landing. GPS can also be used to monitor fish stock. This information can be used for evidence-based policy making to safeguard against overfishing thus improving the sustainability of the industry as a whole.

References


Bornmann, L., Marx, W., Schier, H., Rahm, E., Thor, A. and Daniel, H. D. (2009). Convergent validity of bibliometric Google Scholar data in the field of chemistry - Citation counts for papers that were accepted by Angewandte Chemie International Edition or rejected but published elsewhere, using Google Scholar, Science Citation Index, Scopus, and Chemical Abstracts. Journal of informetrics, 3, pp27-35.


Manyika, J., Chui, M., Brown, B., Bughin, J., Dobbs, R., Roxburgh, C. and Byers, A. H. (2011). *Big data: The next frontier for innovation, competition, and productivity*.


