

## **Big data analytics and enterprises: a bibliometric synthesis of the literature**

Khanra Sayantan, Dhir Amandeep, Mäntymäki Matti

This is a Accepted manuscript (AM) version of a publication  
published by Taylor & Francis  
in Enterprise Information Systems

**DOI:** 10.1080/17517575.2020.1734241

**Copyright of the original publication:** © 2020 Informa UK Limited

### **Please cite the publication as follows:**

Sayantan Khanra, Amandeep Dhir & Matti Mäntymäki (2020) Big data analytics and enterprises: a bibliometric synthesis of the literature, Enterprise Information Systems, DOI: 10.1080/17517575.2020.1734241

This is an Accepted Manuscript of an article published by Taylor & Francis in Enterprise Information Systems on 17 Mar 2020, available online: <http://www.tandfonline.com/10.1080/17517575.2020.1734241>.

**This is a parallel published version of an original publication.  
This version can differ from the original published article.**

## **A bibliometric analysis of the big data analytics literature. Synthesis and framework**

### **Sayantankhanra**

Turku School of Economics, University of Turku, Finland

Indian Institute of Management Rohtak, NH-10 Southern Bypass, Sunaria – 124010,  
Haryana, India.

[sayantankhanra2012@gmail.com](mailto:sayantankhanra2012@gmail.com)

### **\*Amandeep Dhir**

School of Business and Management, LUT University, Lappeenranta, Finland

Turku School of Economics, University of Turku, Finland

Optentia Research Focus Area, North-West University, Vanderbijlpark, South Africa

Email: [amandeep.dhir@lut.fi](mailto:amandeep.dhir@lut.fi)

### **Matti Mäntymäki**

Turku School of Economics, University of Turku, Finland

University of Turku, Finland

[matti.mantymaki@utu.fi](mailto:matti.mantymaki@utu.fi)

\*Corresponding Author

# Big Data Analytics and Enterprises: A Bibliometric Synthesis of the Literature

## Abstract

Understanding the developmental trajectories of big data analytics in the corporate context is highly relevant for information systems research and practice. This study presents a comprehensive bibliometric analysis of applications of big data analytics in enterprises. The sample for this study contained a total of 1727 articles published between 2009 and 2019 as indexed by the Scopus database. The sample represents contributions from 4342 authors, 3515 organizations, 89 countries, and 763 journals on this research topic. The sample was analysed with techniques such as bibliographic coupling, citation analysis, co-word analysis, and co-authorship analysis. Four major thematic areas in the extant literature emerged from the co-citation analysis. The thematic areas were inputs for strategic decision-making, concept development for big data analytics, trends in applications of big data analytics, and efficient supply chain management. The evolution of these thematic areas was documented with dynamic co-citation analysis. The results provide insights into the extant research on big data analytics and can guide future research efforts in this area.

**Keywords:** Big data, analytics, bibliometric analysis, big data analytics, content analysis, predictive analytics, network analysis, prestige analysis

## 1. Introduction

Big data analytics encapsulates various data-intensive approaches capable of analysing large-scale data (Davenport and Patil 2012; Hu et al. 2014). Big data analytics deliver high value in decision-making processes (Rho and Vasilakos 2018; LaValle et al. 2011), primarily because large-scale data offer high granularity, meaning minute data containing in-depth information related to a study objective (George, Haas, and Pentland 2014; McAfee and Brynjolfsson 2012). The extensive use of big data analytics in enterprises may deliver remarkable results (Li et al. 2019), such as the enhancement of firm performance (Akter et al. 2016; Côte-Real, Oliveira, and Ruivo 2017; Popovič et al. 2018), improvements in talent management (Chae 2015; Davenport, Harris, and Shapiro 2010), and better management of supply chains (Gunasekaran et al. 2017; Tan et al. 2015; Zhong et al., 2015). However, enterprises often face several challenges (Jagadish et al. 2014), such as poor data quality (Hazen et al. 2014), concerns for user privacy and data security (Raguseo 2018), and legal and ethical issues (Galloway 2017), among others, while harvesting the benefits of big data analytics.

Research on the use of big data analytics in enterprises has gained significant momentum in recent years (Batistič and van der Laken 2019). Consequently, several reviews have been conducted to summarize the extant knowledge in the field of enterprise information systems. For example, Sheng,

Amankwah-Amoah and Wang (2019) conducted a review of multidisciplinary perspectives on big data analytics from major management domains, like operations research, marketing, and enterprise information systems. Wang, Gunasekaran, Ngai, and Papadopoulos (2016) in turn reviewed the application of big data analytics specific to the domain of supply chain management. Furthermore, recent reviews summarized the niche bodies of knowledge about the applications of big data analytics in the agricultural supply chains (Kamble, Gunasekaran, and Gawankar 2019), and supply chains in disaster management (Akter and Wamba 2019). Prior studies also explored the importance of specific theoretical perspectives, such as resource-based view (Batistič and van der Laken 2019), and dynamic capabilities (Rialti et al. 2019), in exploring the performance enhancement in enterprises by leveraging big data analytics. Despite these important attempts to synthesize the extant literature, knowledge about the application of big data analytics in enterprises appears fragmented (Batistič and van der Laken 2019), and lacks clarity (Sheng, Amankwah-Amoah, and Wang 2019). Hence, there is an evident need for research that provides a comprehensive understanding of the past, present, and future of research on big data analytics in enterprises.

As a result, this study addresses three research questions: (1) what is the current status of research on big data analytics in enterprises? (2) what key thematic areas exist in the extant research, and how have these thematic areas evolved over time? and (3) what potential future research avenues can be identified based on the extant literature? These research questions were answered with analyses of the publications discussing big data analytics in the context of management studies following bibliometric techniques from recent prior research (Caviggioli and Ughetto 2019; Fahimnia, Sarkis, and Davarzani 2015; Xu et al. 2018). These techniques are well recognized for their ability to present structured information about a research topic from a multi-disciplinary perspective through the analysis of a large number of documents (Caviggioli and Ughetto 2019). Bibliometric techniques are established on sound statistical foundations (Xu et al. 2018), leaving little room for the subjective biases that often can affect literature reviews (Durach, Kembro, and Wieland 2017). From these premises, this bibliometric study provides another perspective for organizing the fragmented extant literature of big data analytics in enterprises.

We addressed the first research question with bibliographic coupling, co-occurrence analysis, citation analysis, and co-authorship analysis of prior studies. With respect to the second research question, we employed data clustering from co-citation analysis to identify four major thematic areas in the literature. The first thematic area explores how big data analytics may provide inputs to the managers responsible for strategic decision-making. The second and third thematic areas are dedicated to the concept development and trends in applications of big data analytics, respectively. The fourth thematic area reflects the opportunities and challenges of big data analytics in supply chain

management. Consequently, a dynamic co-citation analysis within each thematic area captures the evolution of thematic areas. With respect to the third research question, we present future research areas in each thematic area. The findings further the understanding of the current state-of-the-art and future development of big data analytics in enterprises.

## 2. Conceptual Background

### 2.1. *Big data analytics: Benefits and challenges*

Big data analytics is conceptualized as a collection of processes that analyse a large volume of data, in structured, semi-structured, or unstructured form, with sophisticated algorithms to obtain precise information about a subject under study (Chaudhuri, Dayal, and Narasayya 2011; Davenport 2013). Big data is often characterized by the five *V*s, namely, volume (amount), velocity (real-time), variety (diversity), value (worth), and veracity (quality) (Kitchin 2014; Wamba et al. 2015). Prior research reports three key benefits of big data analytics in enterprises, as follows:

- (a) *Superior firm practices*. Big data analytics improve practices in different enterprises across industries (Acito and Khatri 2014; Chang, Kauffman, and Kwon 2014). For instance, big data analytics provide value in healthcare management by forecasting the success of medical treatments (Oztekin, Delen, and Kong 2009), making organizational routines more efficient (Wang, Kung, and Byrd 2018), and identifying new avenues of revenue generation (Wang and Hajli 2017).
- (b) *Superior firm strategies*. The predictive power of big data analytics may help enterprises design better strategies (Gandomi and Haider 2015; Yaqoob et al. 2016). For instance, a suitable application of big data analytics may forecast volatility in financial markets more accurately than currently available methods (Sun, Chen, and Yu 2015). The deployment of firm resources in consumer data analytics develops efficient marketing strategies (Erevelles, Fukawa, and Swayne 2016) and helps new products succeed (Xu, Frankwick, and Ramirez 2016).
- (c) *Superior firm performances*. The improved enterprise performance due to big-data-driven practices is also observed in disaster management (Papadopoulos et al. 2017), electronic commerce (Akter and Wamba 2016), the information technology industry (Wamba et al. 2017), and sustainable manufacturing (Dubey et al. 2016) among others. Dubey et al. (2019) studied the capability of big data analytics to improve operations in disaster relief missions through stronger collaborations between civil and military organizations.

Enterprises face several challenges while applying big data analytics in business practices (Jagadish et al. 2014; Raguseo 2018). A lack of transparency regarding the working principle of analytical tools raises ambiguity among the managers who interpret the outputs from big data analytics

(Chen, Chiang, and Storey 2012). The poor quality of input data may yield erroneous output, which is unlikely to be detected by managers who are not well trained to understand the working principles of the software they use (Hazen et al. 2014). Also, human subjectivity may affect core issues in big data analytics, such as input data preparation, algorithm design, and output interpretation among others (Raguseo 2018; Trkman et al. 2010). Moreover, technology-driven correlation may not conform to ethical standards of human logic (Barocas and Selbst 2016), particularly in cases like remote surveillance of people without their permission (Galloway 2017). Thus, privacy and security concerns about user data need to be addressed by organizations so that they continue to benefit from the applications of big data analytics in a sustainable manner (Raguseo 2018; Siemens 2013).

## 2.2. *Bibliometric Protocol*

Bibliometric analysis recognizes key authors, organizations, and countries for their important contributions to the extant literature on a topic (Xu et al. 2018). This methodology is well-recognized for providing a comprehensive summary of a topic of intellectual interest, defining its boundaries, and suggesting agendas for future research with less subjective bias (Caviggioli and Ughetto 2019; Fahimnia, Sarkis, and Davarzani 2015). Prior research has shown that this method has been applied to determine intellectual structures of research topics in multiple domains, such as manufacturing (Caviggioli and Ughetto 2019), sustainability (Fahimnia, Sarkis, and Davarzani 2015), finance (Xu et al. 2018), emerging technologies (Li, Porter, and Suominen 2018), and innovation (van Oorschot, Hofman, and Halman 2018), among others. This study developed a protocol for bibliometric studies from commonly adopted techniques in prior research (Caviggioli and Ughetto, 2019; Fahimnia, Sarkis, and Davarzani 2015; Xu et al., 2018). The protocol begins with defining a topic of intellectual interest, and then follows four sequential steps to provide directions for advancing the research on the topic. The document selection process for a bibliometric analysis involves three sequential tasks for scanning, curating, and reporting a sample profile (Fahimnia, Sarkis, and Davarzani 2015; Ferrieira, 2018). The protocol, as reported in Exhibit A, includes the following techniques:

- (a) *Bibliographic coupling*. This technique of analysing extant literature pairs two publications that refer to a common document (Kessler 1963). The principle behind this technique is that a higher number of shared references between a pair of publications occurs due to their common intellectual capital (Li, Porter, and Suominen 2018). While this principle based on backward citation chaining is commended for providing a succinct assessment of a research topic (van Eck and Waltman 2014), it is often censured for its limited ability to cluster older publications (van Oorschot et al. 2018). Therefore, the results from bibliographic coupling are often

complemented with a technique based on forward citation chaining, such as citation analysis and co-citation analysis (Ferreira 2018; Leung, Sun, and Bai 2017).

- (b) *Citation analysis*. Citation analysis attempts to quantify the level of acceptance a publication has gained in academia in terms of the number of times the publication is cited in other publications (Xu et al. 2018). This technique identifies authors, organizations, and countries contributing popular publications to the extant literature on a research topic. However, citation analysis only takes a publication's popularity among other documents within a sample into consideration, and not its prestige (Ding and Cronin 2011). A publication's prestige, as defined in bibliometric terms, depends on the number of times it is cited by highly regarded publications (Ding and Cronin 2011; Xu et al. 2018). Therefore, the incorporation of prestige analysis into the findings from citation analysis provides a more comprehensive understanding of a publication's impact.
- (c) *Co-citation analysis*. Co-citation analysis is an enhanced measure to identify thematic similarities among publications on a research topic (Ng et al. 2018; Small 1973). Two documents are deemed co-cited when both of them are cited by another publication (Xu et al. 2018). A cluster of co-cited documents form a module of a literature network (Fahimnia, Sarkis, and Davarzani 2015), which is identified from the semantic similarities among publications that independently cite documents from the network (Shiau, Dwivedi, and Yang 2017). Themes emerging from a cluster may represent the key intellectual interests being explored by the documents in the cluster (Shiau et al., 2017; Small 1973). This technique overcomes the limitation of bibliographic coupling by effectively clustering publications extant to a research topic, even the older ones (Ferreira 2018; van Oorschot et al. 2018). Furthermore, a dynamic co-citation analysis provides insights about the evolution of those clusters and consequently, suggests future research directions (Fahimnia, Sarkis, and Davarzani 2015; Xu et al. 2018). Co-citation analysis faces a limitation of its own to represent the contents of cited publications, and the limitation may be addressed with a co-word analysis (Chang, Huang, and Lin 2015).
- (d) *Co-word analysis*. By measuring the degree of co-occurrence of keywords (Callon, Courtial, and Laville 1991), co-word analysis aids researchers to better visualize the core content of publications on a research topic (Leung, Sun, and Bai 2017). Studies have posited that combining co-word analysis and co-citation analysis may provide a profound understanding of the development and intellectual configuration of a research topic, as these two techniques suitably complement each other (Åström 2002; Chang et al. 2015). However, a certain degree

of instability may be inherent in the findings from this analysis due to the evolution in key terms over the years (Leung, Sun, and Bai 2017).

- (e) *Co-authorship analysis*. Co-authorship analysis helps researchers understand collaboration patterns among authors contributing to an area of study from the degree of joint publications (Caviggioli and Ughetto 2019). From the collaborative networks of the authors, connections among organizations and countries affiliated to the authors are also examined (Martínez-López et al. 2018). Authors connected to a collaborative network tend to cite similar publications, and consequently, such networks play a key role within a body of knowledge (Racherla and Hu 2010). Therefore, findings from co-authorship analysis strengthen findings from other bibliometric techniques (Martínez-López et al. 2018). However, this technique is not equipped to recognize the contributions of individual authors (Racherla and Hu 2010).
- (f) *Prestige analysis*. From the number of times a publication is cited, the popularity of the publication may be determined but not its prestige, which is considered another important indicator of influence (Ding and Cronin 2011; Xu et al. 2018). The PageRank algorithm (Brin and Page 1998), originally developed to prioritize web pages appearing in Google search results, may be augmented to identify citation links between papers for the concurrent analysis of both prestige and popularity of a publication (Fahimnia, Sarkis, and Davarzani 2015; Xu et al. 2018). When a publication ( $P_0$ ) is cited by another publication,  $P_i$ , where integer  $i \in [1, N]$ , that is cited  $C(P_i)$  times within a cluster of  $N$  publications, the PageRank of  $P_0$ , or  $PR(P_0)$ , may be computed as follows:

$$PR(P_0) = \frac{(1 - d)}{N} + d \left[ \sum_{i=1}^N \frac{PR(P_i)}{C(P_i)} \right]$$

where  $d$  ( $d \in [0, 1]$ ) represents the damping factor, accounting for a fraction of the random walks that propagate through the citations (Brin and Page 1998; Xu et al. 2018). Co-citation PageRank is regarded as an effective measure for ranking publications according to their influence within a co-citation cluster (Fahimnia, Sarkis, and Davarzani 2015). This augmented PageRank algorithm prioritizes publications that are co-cited with highly co-cited papers following an iterative process from a probability distribution of publications (Brin and Page 1998; Fahimnia, Sarkis, and Davarzani, 2015).

-----  
Insert Exhibit A about here  
-----

### 3. Methodology and descriptive statistics



The Scopus database provides comprehensive coverage of resources for the academic community (Caviggioli and Ughetto 2019; Martínez-López et al. 2018). Therefore, the publications relevant to this study were found from conducting a search with pre-determined terms on the title, keywords, and abstracts on Scopus (Caviggioli and Ughetto 2019; Fahimnia, Sarkis, and Davarzani 2015; Xu et al. 2018).

### *3.1. Scanning phase*

Though the key term for this bibliometric analysis was ‘big data analytics’, it was important to identify other terms that conveyed a similar meaning in the context of the present study. Preliminary exploration on Scopus with the search term ‘big data analytics’, suggested that ‘predictive analytics’ was a common keyword that also represented the same general concept. Consequently, a search using ‘big data analytics OR predictive analytics’ resulted in the appearance of 4879 documents published at any time before 01 August 2019 from the Scopus database. The documents belonged to different categories, such as journal articles, book chapters, and conference papers among others. It may be posited that the combined search term provides an appropriate representation of extant research as search with the individual terms of big data analytics and predictive analytics identified 3309 and 4027 documents, respectively.

### *3.2. Curating phase*

Among the 4879 documents that appeared in the scanning phase, the search results were then restricted to articles published in journals and available in English. They were further refined to focus on the subject areas of business, management, and accounting, and related areas, namely decision sciences, economics, econometrics, finance, energy, environmental science, and social sciences. This refinement process excluded a total of 3152 documents, and the remaining 1727 articles that constituted our sample were exported in research information systems (.ris) and comma-separated values (.csv) formats for further analysis.

### *3.3. Analysing phase*

The first among the 1727 articles were published in 2009, indicating that big data analytics was a fairly new research topic in enterprise research at that time. The number of articles published per year from 2010 to date follows an increasing trend, as revealed by Figure 1. A total of 748 articles were from journals in the areas of business, management, and accounting, whereas 979 articles were published in five related areas, as reported in Figure 2. It was found that a total of 4342 authors affiliated to 3515 organizations and located in 89 countries have contributed in advancing this topic by publishing their research in 763 journals as indexed in Scopus.

---

Insert Figures 1 and 2 about here

---

Within our sample, the most published authors, organizations, and countries based on the number of contributions are listed in Tables 1–3, respectively. Prof. Samuel Fosso Wamba (Toulouse Business School, France), Prof. Rameshwar Dubey (Montpellier Business School, France), and Prof. Shahriar Akter (University of Wollongong, Australia) co-authored the highest number of articles: 17, 14, and 13, respectively. The Toulouse Business School in France (7) had the largest number of published articles among the organizations, followed by the Plymouth Business School of the United Kingdom (6) and the Montpellier Business School in France (5). Also, the United States (712), the United Kingdom (216), and India (153) contributed the highest number of articles on the concerned research topic.

---

Insert Tables 1–3 about here

---

The *Journal of Big Data* (52) was the most prominent publication outlet for research on the application of big data analytics in enterprises in our sample, followed by *Big Data* (38) and *Decision Support Systems* (31). Table 4 reports the top journals ranked according to the total number of articles published within our sample. The dynamic counts of publications in each of these journals are also presented in the same table. From Table 4, it may be noticed that *Decision Support Systems* demonstrated a sustained interest on this topic over the past eight years, whereas the topic found prominence in outlets like the *Journal of Cleaner Production* more recently.

---

Insert Table 4 about here

---

#### **4. Bibliometric analyses of publications**

Bibliometric and network analyses of publications may be conducted using multiple software packages, such as Bibexcel (Fahimnia, Sarkis, and Davarzani 2015); Gephi (Xu et al. 2018), VOSviewer (Caviggioli and Ughetto 2019), Pajek (Persson, Danell, and Schneider 2009) and CiteSpace (Li, Porter, and Suominen 2018). The present study primarily utilized VOSviewer to conduct bibliographic coupling, co-word analysis, co-authorship analysis, and citation analysis. This was because VOSviewer was deemed appropriate software to investigate large datasets and offered a range of sophisticated options that helped to better visualize the results (Fahimnia, Sarkis, and Davarzani 2015; van Eck, and Waltman 2014). Gephi was used for co-citation analysis and prestige analysis due to its specialized capability for dynamic analyses and apt adaptation of the PageRank

algorithm, respectively (Xu et al. 2018). Threshold criteria, when applied, are mentioned with the tables and figures that report findings from the analyses.

This study ranked the key contributors, namely authors, organizations, countries, and journals, identified from the bibliographic-coupling and citation analyses based on the attribute of total link strength. The total link strength for a given item refers to the summation of the number of bibliographic links with other items (van Eck and Waltman 2014). Since the number of co-authors in a publication may influence the total link strength of the publication, the fractional counting of bibliographic coupling links was taken into account throughout this study to adjust for this influence (van Eck and Waltman 2014).

#### *4.1. Bibliographic coupling*

Prior research has performed bibliographic coupling using different units of analyses, such as authors of publications (Fahimnia, Sarkis, and Davarzani 2015; Ferreira 2018), organizations affiliated with the authors (Fahimnia, Sarkis, and Davarzani 2015; Ferreira 2018), countries where organizations were located (Caviggioli and Ughetto 2019; Fahimnia, Sarkis, and Davarzani 2015), and journals from where the publications were sourced (Caviggioli and Ughetto 2019; Xu et al. 2018). The present study acknowledged influential authors, recognized organizations and countries that produced important contributions, and identified important publication outlets with the help of the bibliographic coupling technique (Fahimnia, Sarkis, and Davarzani 2015; Martínez-López et al. 2018). This technique captures the influence of a publication in academia with backward citation chaining (van Eck and Waltman 2014). Tables 5–8 report the top-ranked authors, organizations, countries, and journals, respectively, according to total link strength, as analysed from the bibliographic coupling technique.

-----  
Insert Tables 5–8 about here  
-----

Table 5 reveals that Prof. Samuel Fosso Wamba (Toulouse Business School, France) was the most influential author on the research topic of application of big data analytics in enterprises, followed by Prof. Shahriar Akter (University of Wollongong, Australia), and Prof. Rameshwar Dubey (Montpellier Business School, France). Six of the 10 influential organizations listed in Table 6 were located in Europe, including the top three: the Montpellier Business School (France), Plymouth Business School (United Kingdom), and Toulouse Business School (France). However, the United States was found to be the most influential country publishing on the research topic in English, followed by the United Kingdom and China. Among the publication outlets, the *Journal of Business Research*, *Decision Support Systems*, and the *International Journal of Information Management* were the top three influential journals in the rankings.

#### 4.2. Citation analysis

Findings from bibliographic coupling are often complemented with results from citation analysis (Fahimnia, Sarkis, and Davarzani 2015; Ferrieira 2018) that investigates a publication's popularity based on forward citation chaining among the articles in a sample (Xu et al. 2018). The top authors, organizations, countries, and journals contributing articles as ranked according to the total link strength from the citation analysis are presented in Tables 9–12, respectively. It may be observed that the three most influential authors (Table 5) and organizations (Table 6) identified by the bibliographic coupling technique were also found to be the most popular authors (Table 9) and organizations (Table 10) though in a different order from the citation analysis. Table 11 reports that the United States, United Kingdom, and India produced the most popular contributions to research on the application of big data analytics in enterprises. The *International Journal of Production Economics*, *MIS Quarterly*, and the *International Journal of Information Management* topped the list of popular journals as presented in Table 12.

-----  
Insert Table 9–12 about here  
-----

#### 4.3. Prestige analysis

Influence and popularity of the contributors as reported from the bibliographic coupling and citation analyses, respectively, to a publication may not mirror its prestige (Ding and Cronin 2011). The prestige of publications may be explored using an augmented PageRank algorithm (Xu et al. 2018). Thus, the top prestigious articles from the sample of 1727 articles identified and ranked according to their PageRank Score are presented in Table 13. This table also reports the number of times these articles were cited locally (by other articles in the sample) and globally (beyond the sample). From Table 13, it may be observed that the three most prestigious articles were: Waller and Fawcett (2013); George, Haas, and Pentland (2014); and Wamba et al. (2015). However, the article by Chen, Chiang, and Storey (2012) received the highest number of citations among the top prestigious articles, both locally and globally.

-----  
Insert Table 13 about here  
-----

#### 4.4. Co-word analysis

Intellectual inputs to the conceptualization of a research topic may be derived from analysing the co-occurrence of keywords (Åström, 2002; Fahimnia, Sarkis, and Davarzani 2015), both author keywords and index keywords (van Eck and Waltman, 2014). Authors of the 1727 articles in the sample of the present study suggested 4877 keywords, and the articles were indexed with 5099 keywords. Among

these, the top author keywords and index keywords were identified based on the number of co-occurrences, which was the same as the total link strength in the co-word analysis, and are reported in Tables 14 and 15, respectively. Five keywords ('big data', 'predictive analytics', 'big data analytics', 'data analytics', and 'data mining') appear in both Tables 14 and 15. A study of keywords in the two tables reveals that authors put more emphasis on the data modelling aspects including machine learning, learning analytics, and business analytics, whereas indexers highlighted the strategic use of the emergent models in decision-making, forecasting, and learning systems.

-----  
Insert Tables 14–15 about here  
-----

A total of 59 author keywords and 125 index keywords co-occurred more than 10 times in the sampled articles. Density diagrams of these author keywords and index keywords are respectively presented in Figures 3 and 4, and tentative clusters were identified. Both of the density diagrams have three similar clusters as follows: (a) *Strategic input* with author keywords 'visual analytics', 'decision making', and 'marketing analytics' and index keywords 'analytical framework', 'visualization', and 'performance assessment'. (b) *Structured learning* with author keywords 'learning analytics', 'predictive modelling', 'deep learning' and index keywords 'predictive analytics', 'artificial intelligence', and 'forecasting'. (c) *Demographic targeting* with author keywords 'social media', 'social media analytics', and 'sentiment analysis' and index keywords 'human', 'social media', and 'prediction'. The fourth cluster of author keywords captures the *association between big data analytics and advanced technologies* (author keywords 'Internet of Things', 'Industry 4.0', and 'cloud computing'). However, the fourth cluster of index keywords encompasses aspects of *big data governance* (index keywords 'digital storage', 'distributed computer systems', and 'data handling').

-----  
Insert Figures 3–4 about here  
-----

## 5. Network analyses

### 5.1. Co-authorship analysis

Co-authorship analysis provides critical information about the collaborative networks that shape the intellectual capital of a research topic (Martínez-López et al., 2018). A total of 774 links were found among the co-authors who contributed three or more articles to the sample of this study. Among them, 136 authors who received 10 or more local citations were divided into 21 groups. Four of these groups contained 65 authors and are presented in Figure 5. The most prominent among the 24 authors in the first group with eight, six, and five links, respectively, were Prof. Yingfeng Zhang (Northwestern

Polytechnical University, China), Prof. Han Li (Tianjin University, China), and Prof. Yichuan Wang (University of Sheffield, United Kingdom). The second group of 18 authors included Prof. Yang Liu (Linköping University, Sweden), Prof. Benjamin Hazen (Air Force Institute of Technology, United States), and Prof. Shivam Gupta (Montpellier Business School, France), who were linked with eight, seven, and six authors from different groups, respectively. The third group of 13 authors included Prof. Jiayuan Wang (Shenzhen University, China) with nine links to authors from different groups, followed by Prof. Yanming Sun (South China University of Technology, China) and Dr. Cheng Fan (Shenzhen University, China) with five links each. Among the fourth group of 10 authors, Prof. Angappa Gunasekaran (Charlton College of Business, United States) and Prof. Rameshwar Dubey (Montpellier Business School, France) had 10 links each to authors from different groups, whereas Prof. Samuel Fosso Wamba (Toulouse Business School, France) and Prof. Shahriar Akter (University of Wollongong, Australia) were a part of nine links each.

-----  
Insert Figure 5 about here  
-----

We found a total of 37 connections among organizations that contributed three or more articles to the sample of this study and received 10 or more local citations. Figure 6 highlights these connections that led to the emergence of four groups comprised of 9, 11, 6, and 10 organizations, respectively. The key organizations in the first group were the Montpellier Business School (France), the Hong Kong Polytechnic University (Hong Kong), and Iowa State University (United States). The leading organizations that formed the second group included the Plymouth Business School (United Kingdom), the Toulouse Business School (France), and the Southern University of Science and Technology, Xueyuan (China). The third group contained the University of Wollongong (Australia), the Kent Business School (United Kingdom), and the Neoma Business School (France) among their most connected organizations. In the last group, the Charlton College of Business (United States), the Air Force Institute of Technology (United States), and the Harbin Institute of Technology (China) were prominent.

-----  
Insert Figure 6 about here  
-----

From the co-authorship analysis, we found that there were 34 countries that contributed at least 10 articles to the literature. The collaborative networks among these countries were divided into three groups, as presented in Figure 7. The first group of 15 countries included the likes of the United States, China, and India that were connected to 30, 25, and 18 other countries, respectively. In the second

group of 11 countries, Germany (22), France (20), and Australia (19) stood out in terms of the number of connections. The United Kingdom (29), Italy (20), and the Netherlands (20) were prominent in the third group comprised of eight European countries.

-----  
Insert Figure 7 about here  
-----

### 5.2. *Co-citation analysis*

Co-citation analysis is a powerful technique to identify major themes emerging from the extant literature on a research topic (Fahimnia, Sarkis, and Davarzani 2015; Xu et al. 2018). By following this technique, a total of 164 articles (nodes) were found from the sample of the present study that were co-cited in 722 combinations (edges). Using the modularity tool in Gephi with this set of articles, we identified 14 clusters, four of which include 10 or more nodes. The modularity tool was based on the Louvain algorithm, which maximizes the modularity index by determining an optimal number of partitions based on iterations (Fahimnia, Sarkis, and Davarzani 2015). The first among the four major clusters was comprised of 40 nodes (24.39% of the nodes in the network), followed by the second cluster (35 nodes, 21.34%), third cluster (19 nodes, 11.59%), and fourth cluster (16 nodes, 9.76%). In total, the four large clusters captured 598 edges or 82.83% of the network of co-cited articles. Table 16 reports the thematic area of each cluster and the top three journals in each cluster based on the number of contributions. Furthermore, the 10 most prestigious articles in each cluster as ranked according to their PageRank score are presented in Table 17.

-----  
Insert Table 16–17 about here  
-----

### 5.3 *Dynamic co-citation analysis*

An examination of the dynamic clusters emerging from a co-citation analysis may reveal a profound understanding of the evolution of the literature related to a research topic (Fahimnia, Sarkis, and Davarzani 2015; Xu et al. 2018). Figures 8 and 9 present the evolution of the four major clusters as they developed individually over time. Cluster 1 emerged in 2011 and recently (between 2017 and 2019) demonstrated the greatest growth among the four clusters. This growth signifies that big data applications in strategic decision-making are still an important topic of interest for researchers. Clusters 2 and 3 emerged before the other two clusters (Figure 8.1) and are approaching saturation because no article has been added to Cluster 2 between 2017 and 2019, while only two articles have been added to Cluster 3 in that period (Figure 9). Saturation in the thematic area of concept development for big data analytics suggests that the research topic is now sufficiently mature. Lesser

interest in the third thematic area is possible because big data analytics is now well adopted across industries. The latest trend is to advance applications of big data analytics within specific domains, such as supply chain management, as discussed in the fourth thematic area. Cluster 4 is the youngest among the four clusters (Figure 9) and grew independently of the other clusters (Figure 8.4).

-----  
Insert Figures 8 – 9 about here  
-----

## **6. Content analysis of thematic areas**

### *6.1. Inputs for strategic decision-making*

Articles in this cluster primarily focused on the ways big data analytics may provide inputs to the managers responsible for strategic decision-making. For instance, Akter et al. (2016) developed a hierarchical model that highlights the role of management capability, technology capability, and talent capability in improving firm performance. When data provisioning, analytical capabilities, and people skills meet a certain level of organizational readiness, business firms deliver high-value performances (Popovič et al. 2018). Also, knowledge management with big data analytics makes an information technology firm agile, and hence the firm gains a competitive advantage over its competitors (Côte-Real, Oliveira, and Ruivo 2017). Erevelles, Fukawa, and Swayne (2016) proposed that the deployment of firm resources, such as physical capital, human capital, and organizational capital in consumer data analytics may transform marketing strategies. Furthermore, the success of a new product may depend on the degree of fusion between big data analytics and traditional marketing analytics (Xu et al. 2016).

Recognizing the importance of big data analytics, Waller and Fawcett (2013) called for further research in the domain of supply chain management. A study found that text analytics of microblogs may deliver insights regarding information sharing, recruitments, and informal communication to supply chain managers (Chae 2015). Dubey et al. (2016) constructed and validated a framework to demonstrate the impact of big data analytics in sustainable manufacturing practices. Wang, Kung, and Byrd (2018) suggested that data-driven strategies may enhance the profitability of healthcare organizations. Finally, the theme of this cluster may be summarized by an outlook that big data analytics contribute important insights to decision-making processes that deliver financial value to firms (Sharma, Mithas, and Kankanhalli 2014).

### *6.2. Concept development for big data analytics*

In line with the theme of this cluster, Davenport (2013) documented the evolution of big data analytics in preceding years from a phenomenon that made managers prefer fact-based decisions to a distinctive capability that creates value for business organizations. During the period of evolution, the role of a



data scientist became one of the most sought-after jobs (Davenport and Patil 2012), and big data analytics found application in boosting employee productivity, employee engagement, and retention of key talents for firms (Davenport, Harris, and Shapiro 2010). Acito and Khatri (2014) highlighted that big data analytics delivered value in different industries, such as healthcare, financial accounting, and supply chain management. In supply chain management, big data analytics offer important capabilities for planning, sourcing, and delivering performance using enterprise information systems (Trkman et al. 2010).

McAfee and Brynjolfsson (2012) conceptualized three characteristics of big data, namely, volume, velocity, and variety; while Kitchin (2014) discussed four more characteristics: value, variability, veracity, and visualization. Also, George, Haas, and Pentland (2014) put forward a profound understanding of the nature of big data to set the stage for future work. Jagadish et al. (2014) identified several technical challenges related to the heterogeneity, inconsistency, and incompleteness of big data, as well as the timeliness, privacy concerns, and human bias in analytics. Siemens (2013) pointed out that poor data quality, privacy concerns, and immature cyber laws may pose serious challenges to big data analytics. Furthermore, too much reliance on algorithm-based insights, especially in the context of the application of learning analytics in academics, may damage human innovativeness (Siemens 2013).

### *6.3. Trends in applications of big data analytics*

Yaqoob et al. (2016) analysed the origin of big data from the paradigms of structuralism and functionalism and highlighted the trend in applications of big data analytics over time. Organizations in technology-intensive industries pioneered the process of extracting value from big data analytics (LaValle et al. 2011). However, the decreasing costs of data acquisition spurred the adoption of big data analytics across many industries (Chaudhuri, Dayal, and Narasayya 2011). For instance, the adoption of big data analytics has resulted in remarkable advancements in the healthcare industry in terms of predicting the success rate of medical treatments (Oztekin, Delen, and Kong 2009) and in delivering business value (Wang and Hajli, 2017).

Big data analytics assume a prime importance for the ability to predict human behaviour from a variety of mass data generated in the form of text, audio, and video (Gandomi and Haider 2015). Lycett (2013) articulated how the use of big data analytics by businesses may influence human behaviour in society with an example of a video streaming service. Also, Chang, Kauffman, and Kwon (2014) reported that applications of big data analytics in contexts like online marketplaces and social media might transform human behaviour in society. A systematic literature review summarized business values delivered by big data analytics to electronic commerce (Aktar and Wamba, 2016).

However, Raguseo (2018) cautioned about the uncertain return on investment in deploying technology infrastructure required for harvesting benefits of big data analytics and training employees to use the technologies. The reason for the caution is that employees may show reluctance to organizational changes, and hence, the technology infrastructure may not be effectively integrated into organizational routines (Raguseo 2018).

#### *6.4. Efficient supply chain management*

Reflecting opportunities and challenges of big data analytics in business, an article by Chen et al. (2012) paved the way for emerging research topics in the extant literature. Hezen et al. (2014) introduced the requirement for controlling and monitoring data quality to obtain accurate insights from big data analytics in supply chain management. Hu et al. (2014) developed a systematic framework comprised of four sequential modules for the generation, acquisition, storage, and analytics of big data. To assess the application of big data analytics in supply chain management, a maturity model with four stages, namely functional, process-based, collaborative, agile, and sustainable, was developed by Wang et al. (2016). Gunasekaran et al. (2017) conceptualized the utilization of big data analytics in supply chain management as a process involving three stages, acceptance, routinization, and assimilation.

Zhong et al. (2015) developed a mathematical model based on big data analytics to track logistics trajectory in supply chain management. Another mathematical model demonstrated the importance of big data analytics in developing innovation capabilities in supply chain management (Tan et al. 2015). Wamba et al. (2015) identified a paucity of empirical research that aims to assess the business value delivered by big data analytics. Later, Wamba et al. (2017) empirically validated associations among the capabilities of big data analytics and firm performance. Papadopoulos et al. (2017) also followed an empirical approach to investigate the use of big data analytics to explain resilience in supply chain management.

#### *6.5. Methodologies*

Methodologies followed by the studies in the four clusters provide important insights about the evolution of the clusters. Articles in Cluster 1 provide inputs for strategic decision-making following a variety of methodological approaches such as conceptual (Erevelles, Fukawa, and Swayne 2016; Sharma, Mithas, and Kankanhalli 2014; Waller and Fawcett 2013; Xu et al. 2016), content analysis (Chae, 2015; Wang, Kung, and Byrd 2018), Delphi study (Akter et al. 2016), empirical (Côte-Real, Oliveira, and Ruivo 2017; Dubey et al. 2016), and mixed methods (Popovič et al. 2018). All of the 10 prestigious articles in Cluster 2 were conceptual in nature to better suit the purpose of concept development for big data analytics in enterprises. Cluster 3 featured four conceptual studies (Chang, Kauffman, and Kwon 2014; Chaudhuri, Dayal, and Narasayya 2011; LaValle et al. 2011; Lycett

2013), three systematic literature reviews (Akter and Wamba 2016; Gandomi and Haider 2015; Yaqoob et al. 2016), an empirical study (Raguseo 2018), a qualitative study (Wang and Hajli 2017), and a statistical model (Oztekin, Delen, and Kong 2009). Cluster 4 drew on knowledge for managing a supply chain more efficiently with the help of big data analytics from empirical studies (Wamba et al. 2017; Gunasekaran et al. 2017; Papadopoulos et al. 2017), mathematical models (Tan et al. 2015; Zhong et al. 2015), and a longitudinal case study (Wamba et al. 2015), in addition to conceptual studies (Chen et al. 2012; Hazen et al. 2014; Hu et al. 2014; Wang et al. 2016).

## 7. Discussion

In a quest to provide a comprehensive understanding of the status of research on the applications of big data analytics in enterprises, this study developed a protocol for bibliometric studies and duly addressed three relevant research questions.

### 7.1 Key findings from bibliometric techniques:

The first research question enquired about the current status of research on big data analytics in enterprises. To answer this question, we employed different bibliometric techniques to identify important authors, organizations, countries, journals, collaborations, and studies, as follows:

- (a) *Key authors.* with respect to the influence of individual authors, we recognized the top authors in our sample based on the number of contributions, bibliographic coupling strength, and link strength from the citation analysis in Tables 1, 5, and 9, respectively. The presence of five authors, namely Samuel Fosso Wamba, Rameshwar Dubey, Shahriar Akter, Angappa Gunasekaran, and Stephen Childe, in all three tables suggests that they were among the key contributors to the area according to our criteria.
- (b) *Key organizations.* With respect to the influential organizations, our results indicated that the Toulouse Business School (France), the Plymouth Business School (United Kingdom), and the Montpellier Business School (France) were the organizations with the most published research on big data analytics in enterprises according to our criteria (see Tables 2, 6, and 10).
- (c) *Key countries.* Nine countries, the United States, the United Kingdom, India, China, Australia, France, Germany, Italy, and Hong Kong dominated the research in big data analysis in enterprises according to our criteria (see Tables 3, 7, and 11).
- (d) *Key journals.* We identified the key publication outlets based on the number of contributions (Table 4), bibliographic coupling (Table 8), and citation analysis (Table 12). These were the *International Journal of Information Management*, *Decision Support Systems*, the *Journal of Business Research*, and the *Journal of Cleaner Production*.

- (e) *Key collaborations.* The co-authorship analysis of authors (Figure 5) revealed four key collaborative networks that shaped the extant literature. Also, the collaborative networks of organizations and countries are illustrated in Figures 6 and 7, respectively.
- (f) *Key studies.* We have identified the most prestigious studies examining big data analysis in enterprises (Table 13) over a 10-year period (Figure 1). The respective results imply that the first articles focused on conceptualizing the business value of big data analytics (Chen et al. 2012; LaValle et al. 2011) and the use of big data analytics in enterprises (George, Haas, and Pentland 2014; McAfee and Brynjolfsson 2012) in a general sense. The subsequent highly prestigious studies, in turn, shed light on the value of big data analytics in supply chain management (Hazen et al. 2014; Waller and Fawcett 2013). The more recent articles reported a Delphi study (Akter et al. 2016), developed a maturity model (Wang et al. 2016), and systematically reviewed the literature (Gandomi and Haider 2015; Wamba et al. 2015). However, articles published in the last 2–3 years are yet to gain enough traction to be included in Table 13.

### *7.2 A conceptual framework for big data analytics in enterprises*

The second research question involved key thematic areas emerging from the extant research, and the evolution of those thematic areas. This question was answered by the clustering of the extant literature based on the co-citation analysis and the identification of four major thematic areas (Table 16). The key journals and prestigious articles for the thematic areas are reported in Tables 16 and 17, respectively. The dynamic co-citation analysis demonstrated that the thematic areas related to concept development and trends in applications of big data analytics were approaching saturation. However, the thematic area that provided inputs for strategic decision-making has drawn sustained interest from academia. The results from the dynamic co-citation analysis also imply that efficient supply chain management using big data analytics emerged as the most recent thematic area. In-depth analyses of the contents in each thematic area are potentially valuable additions to the literature related to applications of big data analytics in enterprises. Findings from content analysis are summarized in a conceptual framework for big data analytics in enterprises, as presented in Figure 10.

-----  
**Insert Figure 10 about here**  
-----

### *7.3 Future research directions*

The third research question aimed to identify potential future research avenues from the extant literature. Findings from this study put forward several future research directions in connection with

the conceptual framework (see Figure 10). The future research directions are comprised of the following streams of inquiry:

- (a) *Niche focus on analytical tools.* Big data analytics is a holistic approach for managing, processing, and analysing big data (Wamba et al. 2015). Thus, big data analytics overarches a diverse set of tools, such as data mining (Chen et al. 2012), multimedia analytics (Hu et al. 2014), and cognitive modelling (Siemens 2013). Future research may be dedicated to comprehending the value offered by a specific tool to management practices.
- (b) *Analytics in select management domains.* Value delivered by big data analytics is explored in many allied domains of management including agriculture (Kamble et al. 2019), education (Siemens 2013), health (Wang and Hajli 2017), information technology (Wamba et al. 2017), manufacturing (Dubey et al., 2016), retail (Akter and Wamba 2016), and supply chains (Chae 2015). However, studies exploring applications of big data analytics in domains, such as corporate governance, financial services, hospitality, and sports among others are scarce at present.
- (c) *Enrichment of minor thematic areas.* The findings from the co-citation analysis reported a trend in the literature related to the use of big data analytics within a specific domain. For instance, the intellectual capital of a thematic area was established around the applications of big data analytics in supply chain management. However, literature on the use of big data analytics in other domains, namely education (Siemens 2013), electronic commerce (Chang, Kauffman, and Kwon 2014; Akter and Wamba 2016), healthcare (Oztekin, Delen, and Kong 2009; Wang, Kung, and Byrd 2018, Wang and Hajli 2017) among others appear fragmented.
- (d) *Foundations of empirical studies.* The establishment of a research topic begins with conceptual articles, followed by studies adopting more rigorous methodologies (Fahimnia, Sarkis, and Davarzani 2015). The literature related to the applications of big data analytics in enterprises is transitioning from conceptual studies (Davenport, Harris, and Shapiro 2010; LaValle et al. 2011; McAfee and Brynjolfsson 2012) to empirical studies (Côte-Real, Oliveira, and Ruivo 2017; Papadopoulos et al. 2017; Wamba et al. 2017). Therefore, the development of measurement instruments to empirically test benefits offered by big data analytics to management practices may significantly contribute to the extant literature (Gunasekaran et al. 2017).
- (e) *Legal and ethical issues.* The benefits of big data analytics in enterprises may be challenged by concerns for user privacy and data security (Raguseo 2018; Siemens 2013). However, existing corporate laws are not sufficient for delivering justice in disputes in relation to such challenges (Galloway 2017). Therefore, researchers in corporate governance and business law may need to

dedicate research to develop legal frameworks that would sustain the value offered by big data analytics in enterprises while maintaining sound ethical practices.

#### *7.4 Theoretical contributions*

This comprehensive bibliometric study may act as a ‘one-stop shop’ for the recent literature on big data analytics in enterprises. By acknowledging the key contributors to the extant literature, this study reveals important insights into the composition of the extant literature. This study structured the extant literature by identifying key thematic areas of research on big data analytics in enterprises. The evolution of the thematic areas was documented by their dynamic research focus in the extant literature. The thematic areas serve as the building blocks of the conceptual framework developed in this study. Actionable future research directions emerging from the conceptual framework will appeal to interested researchers. Future researchers may also adopt the protocol for bibliometric studies developed in this study to provide structure to other fragmented bodies of knowledge.

#### *7.5 Practical contributions*

Many enterprises fail to gain a desired return on investment in acquiring resources required for harvesting the benefits of big data analytics. This study may help such enterprises to conduct better cost-benefit analyses by providing a comprehensive overview of the potential applications of big data analytics. Enterprises often enjoy the benefits of big data analytics at the expense of users’ data privacy. Policymakers may consider developing legal frameworks that would prevent the exploitation of user data by enterprises. Ethical enterprises may volunteer to address concerns regarding users’ data privacy. From a methodological standpoint, research and development teams in non-academic organizations may find the protocol for bibliometric studies helpful in selecting documents suitable to serve their purposes.

### **8. Conclusion**

The present study addresses the paucity of research about the evolution and current status of big data analytics in enterprises. However, the findings from the present study are subject to the inherent limitations of our sample, and the software that analysed the sample. The sample of the study was constituted of journal articles in English from the Scopus database. Future studies may address these limitations by exploring different databases, including book chapters and conference proceedings, and incorporating documents in multiple languages. Also, VOSviewer software does not distinguish the contribution of the first author from that of other authors to an article while ranking key contributors in the extant literature. Nevertheless, the present study provides structure to the fragmented literature on big data analytics in enterprises by offering one of the earliest bibliometric studies on the extant literature. This study analysed important thematic areas related to big data analytics in enterprises and

proposed actionable future research agendas to advance the extant literature. The thematic areas and future research agendas are presented in a conceptual framework for big data analytics in enterprises. Finally, the protocol for bibliometric studies developed in this study may appeal to future researchers interested in contributing a comprehensive overview of a research topic.

## References

- Acito, F., and V., Khatri. 2014. "Business analytics: Why now and what next?" *Business Horizons* 57: 565-570.
- Akter, S., & Wamba, S. F. (2019). Big data and disaster management: a systematic review and agenda for future research. *Annals of Operations Research*, 283(1-2), 939-959.
- Akter, S., and S. F., Wamba, 2016. "Big data analytics in E-commerce: a systematic review and agenda for future research." *Electronic Markets* 26(2): 173-194.
- Akter, S., S. F., Wamba, A., Gunasekaran, R., Dubey, and S. J., Childe. 2016. "How to improve firm performance using big data analytics capability and business strategy alignment?" *International Journal of Production Economics* 182: 113-131.
- Araz, O. M., D., Olson, and A., Ramirez-Nafarrate. 2019. "Predictive analytics for hospital admissions from the emergency department using triage information" *International Journal of Production Economics* 208: 199-207.
- Åström, F. 2002. "Visualizing library and information science concept spaces through keyword and citation based maps and clusters." Paper presented at the fourth international conference on conceptions of Library and Information Science, July.
- Barocas, S., and A. D., Selbst. 2016. "Big data's disparate impact." *California Law Review* 104(3): 671-732.
- Batistič, S., & van der Laken, P. (2019). History, evolution and future of big data and analytics: a bibliometric analysis of its relationship to performance in organizations. *British Journal of Management*, 30(2), 229-251.
- Brin, S., and L., Page. 1998. "The anatomy of a large-scale hypertextual web search engine." *Computer Networks and ISDN Systems* 30(1-7): 107-117.
- Callon, M., J. P., Courtial and F., Laville. 1991. "Co-word analysis as a tool for describing the network of interactions between basic and technological research: The case of polymer chemistry." *Scientometrics* 22(1): 155-205.
- Caviggioli, F., and E., Ughetto. 2019. "A bibliometric analysis of the research dealing with the impact of additive manufacturing on industry, business and society." *International Journal of Production Economics* 208: 254-268.
- Chae, B. K. 2015. "Insights from hashtag# supply chain and Twitter Analytics: Considering Twitter and Twitter data for supply chain practice and research." *International Journal of Production Economics* 165: 247-259.
- Chang, R. M., R. J., Kauffman, and Y., Kwon. 2014. "Understanding the paradigm shift to computational social science in the presence of big data." *Decision Support Systems* 63: 67-80.
- Chang, Y. W., M. H., Huang, and C. W., Lin. 2015. "Evolution of research subjects in library and information science based on keyword, bibliographical coupling, and co-citation analyses." *Scientometrics* 105(3): 2071-2087.
- Chaudhuri, S., U., Dayal, and V., Narasayya. 2011. "An overview of business intelligence technology." *Communications of the ACM* 54(8): 88-98.
- Chen, C. P., and C. Y., Zhang. 2014. "Data-intensive applications, challenges, techniques and technologies: A survey on Big Data". *Information Sciences* 275: 314-347.
- Chen, H., R. H., Chiang, and V. C., Storey. 2012. "Business intelligence and analytics: From big data to big impact." *MIS Quarterly* 36(4): 1165-1188.

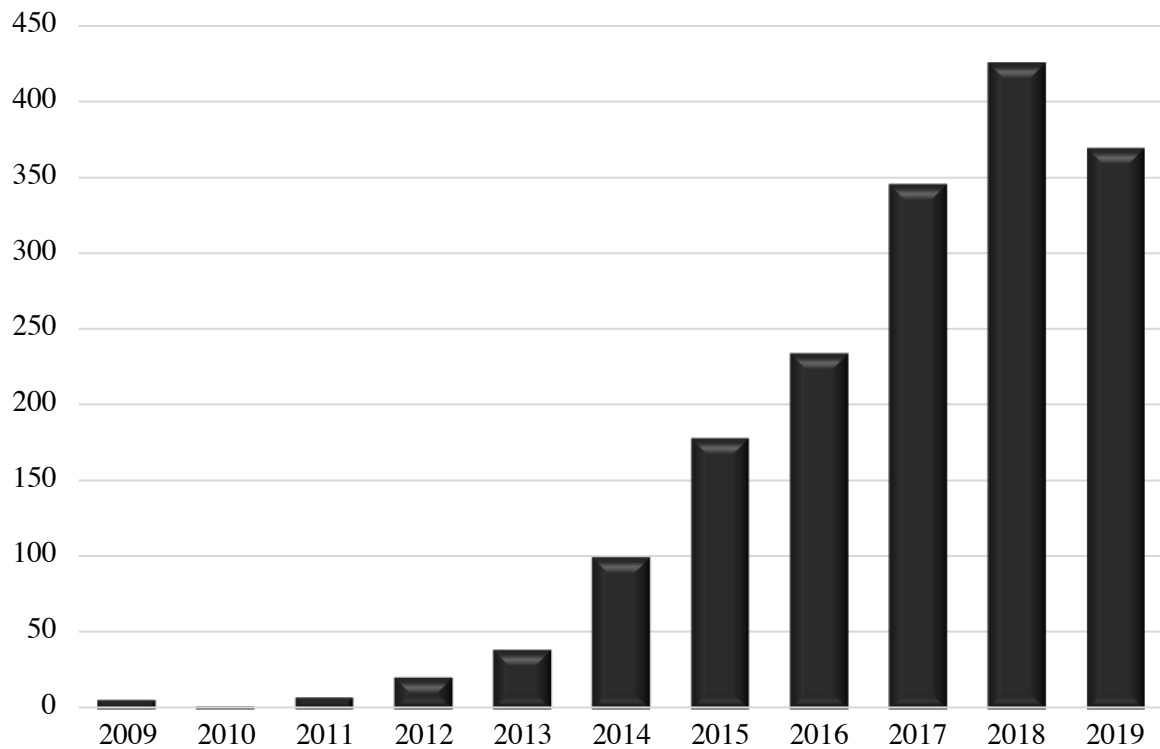
- Côrte-Real, N., T., Oliveira, and P., Ruivo. 2017. "Assessing business value of Big Data Analytics in European firms." *Journal of Business Research* 70: 379-390.
- Davenport, T. H. 2013. "Analytics 3.0." *Harvard Business Review* 91(12): 64-72.
- Davenport, T. H., and D. J. Patil, 2012. "Data scientist: The Sexiest Job of the 21st Century." *Harvard Business Review*, 90(5): 70-76.
- Davenport, T. H., J., Harris, and J., Shapiro. 2010. "Competing on talent analytics." *Harvard Business Review* 88(10): 52-58.
- Ding, Y., and B. Cronin. 2011. "Popular and/or prestigious? Measures of scholarly esteem." *Information Processing and Management* 47(1): 80-96.
- Dubey, R., A., Gunasekaran, S. J., Childe, D., Roubaud, S. F., Wamba, M., Giannakis, and C., Foropon. 2019. "Big data analytics and organizational culture as complements to swift trust and collaborative performance in the humanitarian supply chain." *International Journal of Production Economics* 210: 120-136.
- Dubey, R., A., Gunasekaran, S. J., Childe, S. F., Wamba, and T., Papadopoulos. 2016. "The impact of big data on world-class sustainable manufacturing." *The International Journal of Advanced Manufacturing Technology* 84(1-4): 631-645.
- Durach, C. F., J., Kembro, and A., Wieland. 2017. "A new paradigm for systematic literature reviews in supply chain management." *Journal of Supply Chain Management* 53(4): 67-85.
- Erevelles, S., N., Fukawa, and L., Swayne. 2016. "Big Data consumer analytics and the transformation of marketing." *Journal of Business Research* 69(2): 897-904.
- Fahimnia, B., J., Sarkis, and H. Davarzani. 2015. "Green supply chain management: A review and bibliometric analysis." *International Journal of Production Economics* 162: 101-114.
- Ferreira, F. A. 2018. "Mapping the field of arts-based management: Bibliographic coupling and co-citation analyses." *Journal of Business Research* 85: 348-357.
- Galloway, K. 2017. "Big Data: A case study of disruption and government power." *Alternative Law Journal* 42(2): 89-95.
- Gandomi, A., and M. Haider. 2015. "Beyond the hype: Big data concepts, methods, and analytics." *International Journal of Information Management* 35(2): 137-144.
- George, G., M. R., Haas, and A. Pentland. 2014. "Big data and management." *Academy of Management Journal*, 57(2): 321-326.
- Gunasekaran, A., T., Papadopoulos, R., Dubey, S. F., Wamba, S. J., Childe, B., Hazen, and S., Akter. 2017. "Big data and predictive analytics for supply chain and organizational performance." *Journal of Business Research* 70: 308-317.
- Harerimana, G., B., Jang, J. W., Kim, and H. K., Park. 2018. "Health big data analytics: A technology survey." *IEEE Access* 6: 65661-65678.
- Hashem, I. A. T., I., Yaqoob, N. B., Anuar, S., Mokhtar, A., Gani, and S. U., Khan. 2015. "The rise of "big data" on cloud computing: Review and open research issues." *Information Systems* 47: 98-115.
- Hazen, B. T., C. A., Boone, J. D., Ezell, and L. A., Jones-Farmer. 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: 72-80.
- Hu, H., Y., Wen, T. S., Chua, and X., Li. 2014. "Toward scalable systems for big data analytics: A technology tutorial." *IEEE Access* 2: 652-687.
- Jagadish, H. V., J., Gehrke, A., Labrinidis, Y., Papakonstantinou, J. M., Patel, R., Ramakrishnan, and C., Shahabi. 2014. "Big data and its technical challenges." *Communications of the ACM* 57(7): 86-94.
- Kamble, S. S., A., Gunasekaran, and S. A., Gawankar. 2019. "Achieving sustainable performance in a data-driven agriculture supply chain: A review for research and applications." *International Journal of Production Economics* 219:179-184.
- Kessler, M. M. 1963. "Bibliographic coupling between scientific papers." *American Documentation* 14(1): 10-25.



- Kitchin, R. 2014. "Big Data, new epistemologies and paradigm shifts." *Big data and society* 1(1): 1-12.
- LaValle, S., E., Lesser, R., Shockley, M. S., Hopkins, and N., Kruschwitz. 2011. "Big data, analytics and the path from insights to value." *MIT Sloan Management Review* 52(2): 21-32.
- Leung, X. Y., J., Sun, and B., Bai. 2017. "Bibliometrics of social media research: A co-citation and co-word analysis." *International Journal of Hospitality Management* 66: 35-45.
- Li, G., J., Tan, and S. S., Chaudhry. 2019. "Industry 4.0 and big data innovations." *Enterprise Information Systems* 13(2): 145-147.
- Li, M., A. L., Porter, and A., Suominen. 2018. "Insights into relationships between disruptive technology/innovation and emerging technology: A bibliometric perspective." *Technological Forecasting and Social Change* 129: 285-296.
- Lycett, M. 2013. "'Datafication': Making sense of (big) data in a complex world." *European Journal of Information Systems* 22: 381-386.
- Martínez-López, F. J., J. M., Merigó, L., Valenzuela-Fernández and C., Nicolás. 2018. "Fifty years of the European Journal of Marketing: a bibliometric analysis." *European Journal of Marketing* 52(1/2): 439-468.
- McAfee, A., and E., Brynjolfsson. 2012. "Big data: the management revolution." *Harvard Business Review* 90(10): 1-9.
- Ng, C. K., C. H., Wu, K. L., Yung, W. H., Ip, and T. Cheung. 2018. "A semantic similarity analysis of Internet of Things." *Enterprise Information Systems* 12(7): 820-855.
- Oztekin, A., D., Delen, and Z. J. Kong. 2009. "Predicting the graft survival for heart-lung transplantation patients: An integrated data mining methodology." *International Journal of Medical Informatics* 78(12): e84-e96.
- Papadopoulos, T., A., Gunasekaran, R., Dubey, N., Altay, S. J., Childe, and S. F., Wamba. 2017. "The role of Big Data in explaining disaster resilience in supply chains for sustainability." *Journal of Cleaner Production* 142: 1108-1118.
- Persson, O., R., Danell, and J. W. Schneider. 2009. "How to use Bibexcel for various types of bibliometric analysis." *Celebrating scholarly communication studies: A Festschrift for Olle Persson at his 60th Birthday* 5: 9-24.
- Popovič, A., R., Hackney, R., Tassabehji, and M., Castelli. 2018. "The impact of big data analytics on firms' high value business performance." *Information Systems Frontiers* 20(2): 209-222.
- Racherla, P., and C., Hu. 2010. "A social network perspective of tourism research collaborations." *Annals of Tourism Research* 37(4): 1012-1034.
- Raguseo, E. 2018. "Big data technologies: An empirical investigation on their adoption, benefits and risks for companies." *International Journal of Information Management* 38(1): 187-195.
- Rho, S., and A. V., Vasilakos. 2018. "Intelligent collaborative system and service in value network for enterprise computing." *Enterprise Information Systems* 12(1): 1-3.
- Rialti, R., Marzi, G., Ciappei, C., & Busso, D. (2019). Big data and dynamic capabilities: a bibliometric analysis and systematic literature review. *Management Decision*.
- Sharma, R., S., Mithas, and A., Kankanhalli. 2014. "Transforming decision-making processes: a research agenda for understanding the impact of business analytics on organizations." *European Journal of Information Systems* 23(4): 433-441.
- Sheng, J., J., Amankwah-Amoah, and X., Wang. 2017. "Technology in the 21st century: new challenges and opportunities." *Technological Forecasting and Social Change*, 143, 321-335.
- Shiau, W. L., Y. K., Dwivedi, and H. S., Yang. 2017. "Co-citation and cluster analyses of extant literature on social networks." *International Journal of Information Management* 37(5): 390-399.
- Siemens, G. 2013. "Learning analytics: The emergence of a discipline." *American Behavioral Scientist* 57(10): 1380-1400.

- Small, H. 1973. "Co-citation in the scientific literature: A new measure of the relationship between two documents." *Journal of the American Society for information Science* 24(4): 265-269.
- Sun, E. W., Y. T., Chen, and M. T., Yu. 2015. "Generalized optimal wavelet decomposing algorithm for big financial data." *International Journal of Production Economics* 165: 194-214.
- Tan, K. H., Y., Zhan, G., Ji, F., Ye, and C., Chang. 2015. "Harvesting big data to enhance supply chain innovation capabilities: An analytic infrastructure based on deduction graph." *International Journal of Production Economics* 165: 223-233.
- Trieu, V. H. 2017. "Getting value from Business Intelligence systems: A review and research agenda." *Decision Support Systems* 93: 111-124.
- Trkman, P., K., McCormack, M. P. V., De Oliveira, and M. B., Ladeira. 2010. "The impact of business analytics on supply chain performance." *Decision Support Systems* 49(3): 318-327.
- Van Eck, N. J., and L., Waltman. 2014. "Visualizing bibliometric networks." In *Measuring Scholarly Impact*, 285-320. Cham: Springer.
- Van Oorschot, J. A., E., Hofman, and J. I., Halman. 2018. "A bibliometric review of the innovation adoption literature." *Technological Forecasting and Social Change* 134: 1-21.
- Waller, M. A., and S. E., Fawcett. 2013. "Data science, predictive analytics, and big data: a revolution that will transform supply chain design and management." *Journal of Business Logistics* 34(2): 77-84.
- Wamba, S. F., A., Gunasekaran, S., Akter, S. J. F., Ren, R., Dubey, and S. J., Childe. 2017. "Big data analytics and firm performance: Effects of dynamic capabilities." *Journal of Business Research* 70: 356-365.
- Wamba, S. F., S., Akter, A., Edwards, G., Chopin, and D., Gnanzou. 2015. "How 'big data' can make big impact: Findings from a systematic review and a longitudinal case study." *International Journal of Production Economics* 165: 234-246.
- Wang, G., Gunasekaran, A., Ngai, E. W., and T., Papadopoulos. 2016. "Big data analytics in logistics and supply chain management: Certain investigations for research and applications." *International Journal of Production Economics* 176: 98-110.
- Wang, Y., and N., Hajli. 2017. "Exploring the path to big data analytics success in healthcare." *Journal of Business Research* 70: 287-299.
- Wang, Y., L., Kung, and T. A. Byrd. 2018. "Big data analytics: Understanding its capabilities and potential benefits for healthcare organizations." *Technological Forecasting and Social Change* 126: 3-13.
- Xu, X., X., Chen, F., Jia, S., Brown, Y., Gong, and Y., Xu. 2018. "Supply chain finance: A systematic literature review and bibliometric analysis." *International Journal of Production Economics* 204: 160-173.
- Xu, Z., G. L., Frankwick, and E., Ramirez. 2016. "Effects of big data analytics and traditional marketing analytics on new product success: A knowledge fusion perspective." *Journal of Business Research* 69(5): 1562-1566.
- Yaqoob, I., I. A. T., Hashem, A., Gani, S., Mokhtar, E., Ahmed, N. B., Anuar, and A. V., Vasilakos. 2016. "Big data: From beginning to future." *International Journal of Information Management* 36(6): 1231-1247.
- Zhong, R. Y., G. Q., Huang, S., Lan, Q. Y., Dai, X., Chen, and T., Zhang. 2015. "A big data approach for logistics trajectory discovery from RFID-enabled production data." *International Journal of Production Economics* 165: 260-272.

**Figure 1: Articles published per year**

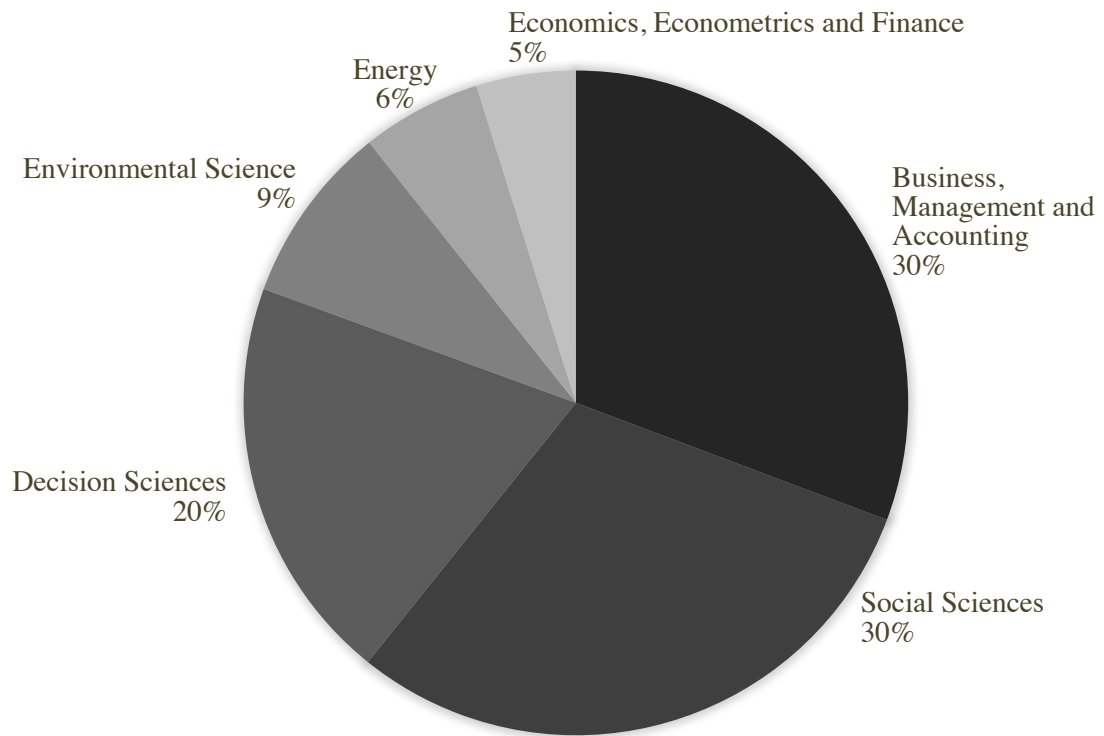


[Data source<sup>1</sup>: Scopus; search date: August 01, 2019]

<sup>1</sup>Scopus search string:

(TITLE-ABS-KEY (big AND data AND analytics) OR TITLE-ABS-KEY (predictive AND analytics)) AND DOCTYPE (ar) AND (LIMIT-TO ( SRCTYPE, "j")) AND (LIMIT-TO (SUBJAREA, "BUSI") OR LIMIT-TO ( SUBJAREA, "SOCI" ) OR LIMIT-TO (SUBJAREA, "DECI") OR LIMIT-TO ( SUBJAREA, "ENVI" ) OR LIMIT-TO (SUBJAREA, "ENER") OR LIMIT-TO ( SUBJAREA, "ECON")) AND (LIMIT-TO (LANGUAGE, "English"))

**Figure 2: Contribution by subject area**



[Data source: Scopus; search date: August 01, 2019]

**Table 1: Top 10 authors based on the number of contributions**

<b>Author</b>	<b>Articles</b>
Wamba, S.F.	17
Dubey, R.	14
Akter, S.	13
Gunasekaran, A.	12
Wang, Y.	11
Wang, L.	11
Zhang, Y.	10
Childe, S.J.	9
Liu, Y.	9
Delen, D.	9

[Data source: Scopus; search date: August 01, 2019]

**Table 2: Top 10 organisations based on the number of contributions**

<b>Organisation*</b>	<b>Articles</b>
Toulouse Business School, Toulouse University, France	7
Plymouth Business School, Plymouth University, United Kingdom	6
Montpellier Business School, Montpellier Research in Management, France	5
University of Wollongong, New South Wales, Australia	5
Department of Engineering Technology, Mississippi Valley State University, United States	4
Neoma Business School, Rouen, France	4
Charlton College of Business University of Massachusetts, Dartmouth, United States	3
City University of Hong Kong, Kowloon Tong, Hong Kong	3
College of Business, Abu Dhabi University, United Arab Emirates	3
Department of Management and Engineering, Linköping University, Sweden	3

\* Affiliation of the first author of an article

[Data source: Scopus; search date: August 01, 2019]

**Table 3: Top 10 countries based on the number of contributions**

<b>Country*</b>	<b>Articles</b>
United States	712
United Kingdom	216
India	153
China	151
Australia	121
Germany	79
Italy	60
France	57
Canada	56
Hong Kong	41

\* Location of an organization affiliated to the first author of an article

[Data source: Scopus; search date: August 01, 2019]

**Table 4: Top 10 journals based on the number of contributions**

<b>Journal title</b>	<b>2012</b>	<b>2013</b>	<b>2014</b>	<b>2015</b>	<b>2016</b>	<b>2017</b>	<b>2018</b>	<b>2019*</b>	<b>Total*</b>
Journal of Big Data	0	0	1	12	5	12	7	15	52
Big Data	0	5	3	14	7	6	2	1	38
Decision Support Systems	1	1	3	2	5	9	7	3	31
International Journal of Recent Technology and Engineering	0	0	0	0	0	0	1	29	30
Sustainability Switzerland	0	0	0	0	1	7	12	8	28
International Journal of Information Management	0	0	1	2	4	3	6	7	23
Journal of Cleaner Production	0	0	1	0	0	5	7	10	23
Big Data Research	0	0	1	4	6	2	7	2	22
European Journal of Operational Research	0	0	0	3	4	4	5	4	20
Journal of Business Research	0	0	0	1	4	8	1	5	19

\* as of August 01, 2019

[Data source: Scopus; search date: August 01, 2019]



**Table 5: Top 10 authors from the bibliographic coupling**

<b>Author</b>	<b>Total link strength</b>
Wamba, S.F.	1485.65
Akter, S.	1266.92
Dubey, R.	1242.40
Gunasekaran, A.	1001.92
Childe, S.J.	963.18
Zhang, Y.	812.41
Liu, Y.	785.65
Papadopoulos, T.	776.34
Chen, H.	756.05
Hazen, B.T.	703.68

**Table 6: Top 10 organizations from the bibliographic coupling**

<b>Organisation</b>	<b>Total link strength</b>
Montpellier Business School, Montpellier Research in Management, France	703.75
Plymouth Business School, Plymouth University, United Kingdom	663.37
Toulouse Business School, Toulouse University, France	493.01
Department of Management and Engineering, Linköping University, Sweden	484.41
School of Business and Public Administration, California State University, United States	413.46
Key Laboratory of Contemporary Design and Integrated Manufacturing Technology in Shaanxi, Ministry of Education, China	399.34
Department of Production, University of Vaasa, Finland	393.44
Research and Development Institute in Shenzhen, Northwestern Polytechnical University, China	364.91
Kent Business School, University of Kent, United Kingdom	357.46
Department of Operational Sciences, Air Force Institute of Technology, United States	353.04

**Table 7: Top 10 countries from the bibliographic coupling**

<b>Country</b>	<b>Total link strength</b>
United States	14064.81
United Kingdom	8380.72
China	5850.74
Australia	4005.91
India	3565.14
France	2968.53
Germany	2543.53
Italy	2057.97
Hong Kong	1711.96
Taiwan	1656.28

**Table 8: Top 10 journals from the bibliographic coupling**

<b>Journal</b>	<b>Total link strength</b>
Journal of Business Research	561.22
Decision Support Systems	535.80
International Journal of Information Management	531.91
Technological Forecasting and Social Change	448.54
International Journal of Production Economics	439.89
British Journal of Management	428.65
Journal of Big Data	409.84
Journal of Cleaner Production	405.56
International Journal of Logistics Management	397.05
European Journal of Operational Research	388.27

**Table 9: Top 10 authors from citation analysis**

<b>Author</b>	<b>Total link strength</b>
Wamba, S.F.	851
Dubey, R.	736
Akter, S.	714
Gunasekaran, A.	661
Chen, H.	606
Chiang, R.H.L.	575
Storey, V.C.	563
Childe, S.J.	551
Papadopoulos, T.	493
Waller, M.A.	488

**Table 10: Top 10 organizations from citation analysis**

<b>Organisation</b>	<b>Total link strength</b>
Toulouse Business School, Toulouse University, France	430
Plymouth Business School, Plymouth University, United Kingdom	396
Montpellier Business School, Montpellier Research in Management, France	323
Carl H. Lindner College of Business, University of Cincinnati, United States	229
Eller College of Management, University of Arizona, United States	229
J. Mack Robinson College of Business, Georgia State University, United States	229
Business Administration, Weber State University, Arkansas, United States	218
Charlton College of Business University of Massachusetts, Dartmouth, United States	209
Department of Management and Engineering, Linköping University, Sweden	201
Department of Marketing and Supply Chain Management, College of Business Administration, University of Tennessee, United States	196

**Table 11: Top 10 countries from citation analysis**

<b>Country</b>	<b>Total link strength</b>
United States	2130
United Kingdom	1197
India	733
China	702
France	680
Australia	613
Canada	292
Italy	289
Germany	234
Hong Kong	167

**Table 12: Top 10 journals from citation analysis**

<b>Journal</b>	<b>Total link strength</b>
International Journal of Production Economics	208
MIS Quarterly: Management Information Systems	197
International Journal of Information Management	195
Journal of Business Research	193
Journal of Cleaner Production	125
Information and Management	105
Decision Support Systems	86
Business Process Management Journal	84
British Journal of Management	82
Technological Forecasting and Social Change	76



**Table 13: Top 10 articles from prestige analysis**

<b>Article</b>	<b>PageRank Score</b>	<b>Local citation count*</b>	<b>Global citation count<sup>#</sup></b>
Waller and Fawcett (2013)	0.018236	47	715
George, Haas, and Pentland (2014)	0.017412	28	647
Wamba et al. (2015)	0.016653	31	678
LaValle, et al. (2011)	0.016494	23	1324
Hazen et al. (2014)	0.015785	32	416
McAfee and Brynjolfsson (2012)	0.015055	34	3599
Chen, Chiang, and Storey (2012)	0.014683	114	4052
Wang et al. (2016)	0.014237	11	364
Gandomi and Haider (2015)	0.013637	65	1780
Akter et al. (2016)	0.012741	23	201

\* cited by articles in the sample

<sup>#</sup> Source: Google scholar (Date: August 10, 2019)

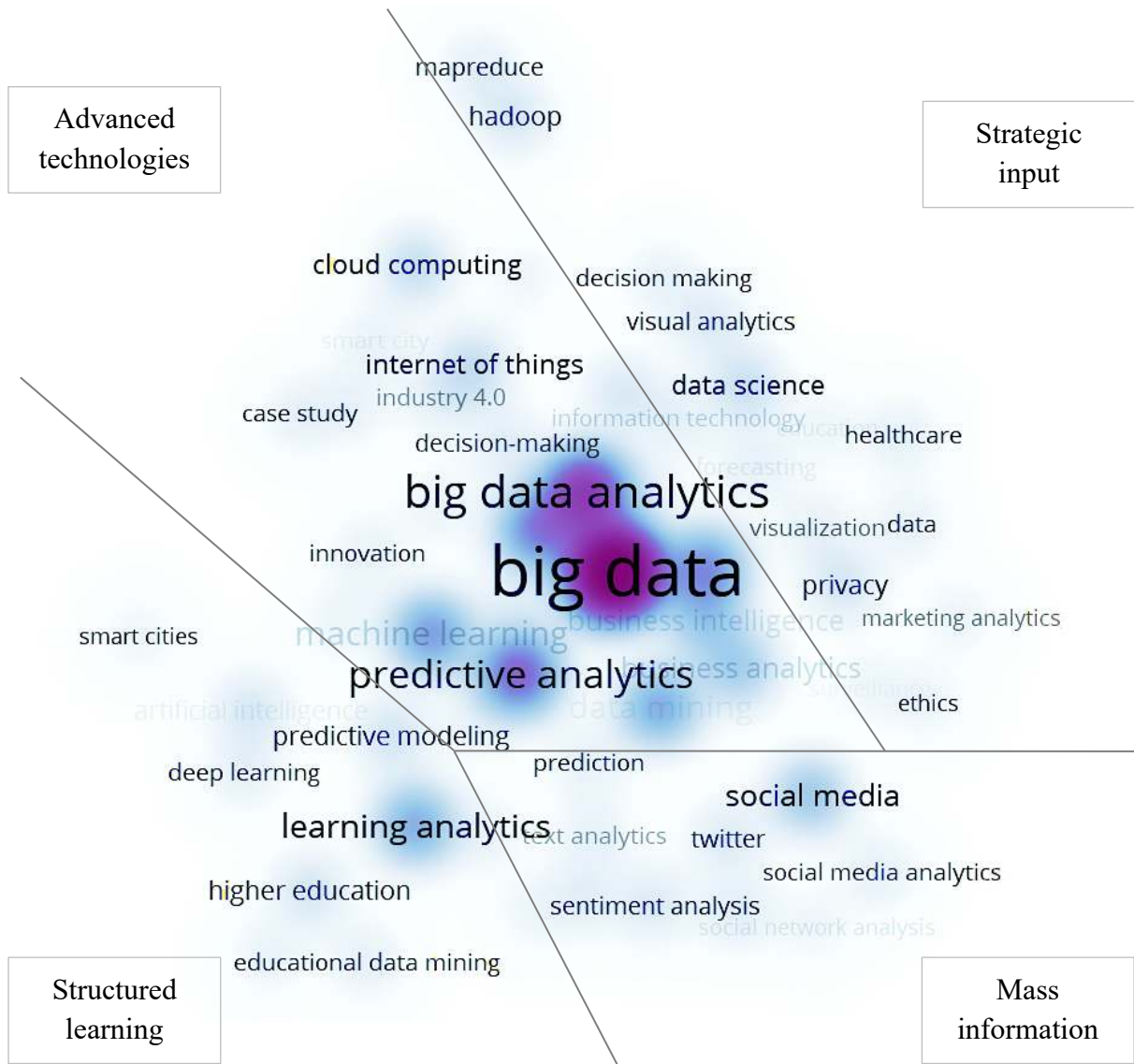
**Table 14: Top 10 author keywords from co-word analysis**

<b>Keyword</b>	<b>Co-occurrences</b>	<b>Total link strength</b>
Big data	590	590
Big data analytics	222	222
Predictive analytics	158	158
Data analytics	124	124
Analytics	114	114
Machine learning	105	105
Data mining	92	92
Learning analytics	91	91
Business analytics	63	63
Business intelligence	56	56

**Table 15: Top 10 index keywords from co-word analysis**

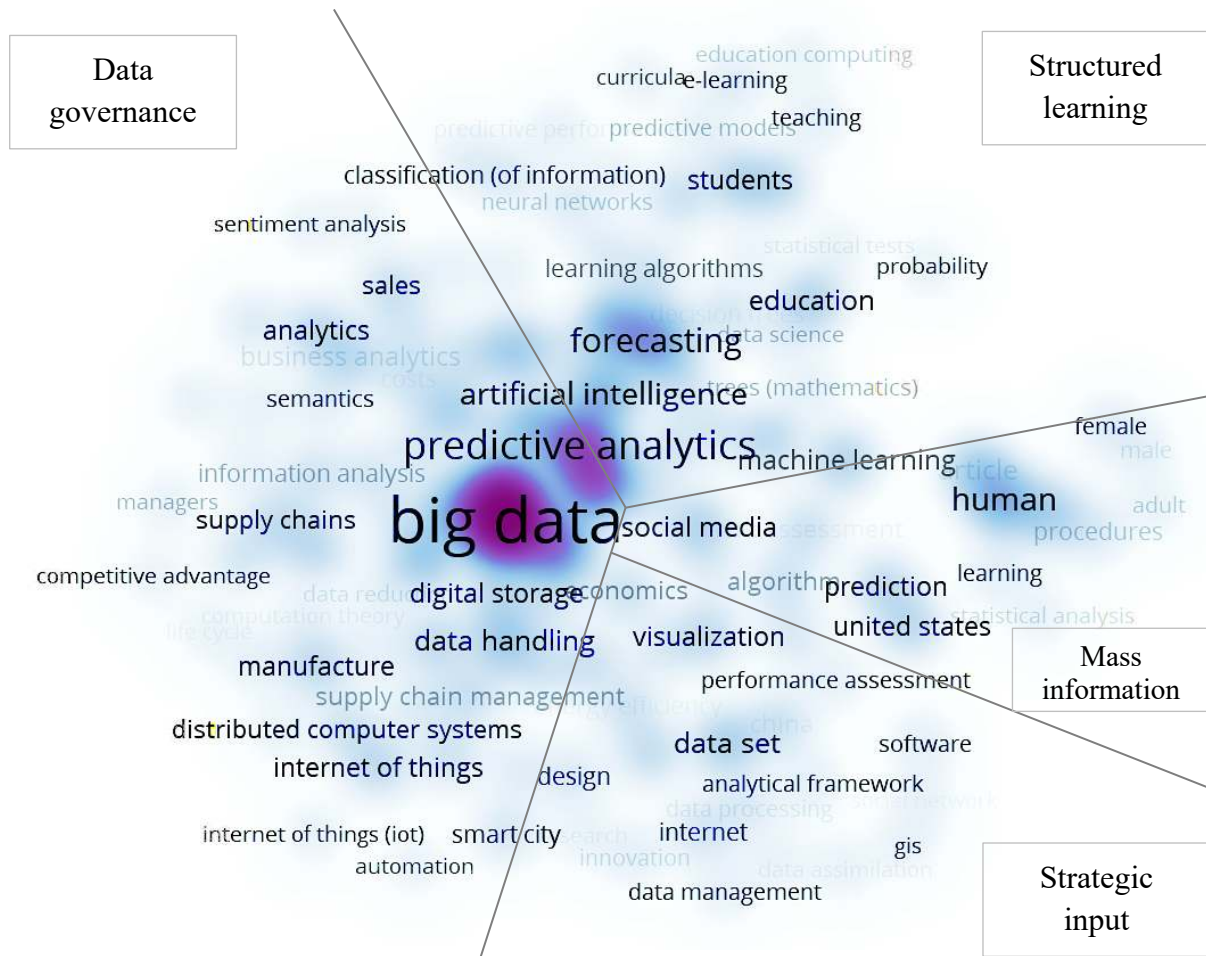
<b>Keyword</b>	<b>Co-occurrences</b>	<b>Total link strength</b>
Big data	370	370
Data analytics	135	135
Predictive analytics	123	123
Data mining	122	122
Decision making	107	107
Forecasting	76	76
Learning systems	68	68
Information management	62	62
Artificial intelligence	61	61
Big data analytics	59	59

**Figure 3: Density diagram of author keywords\***

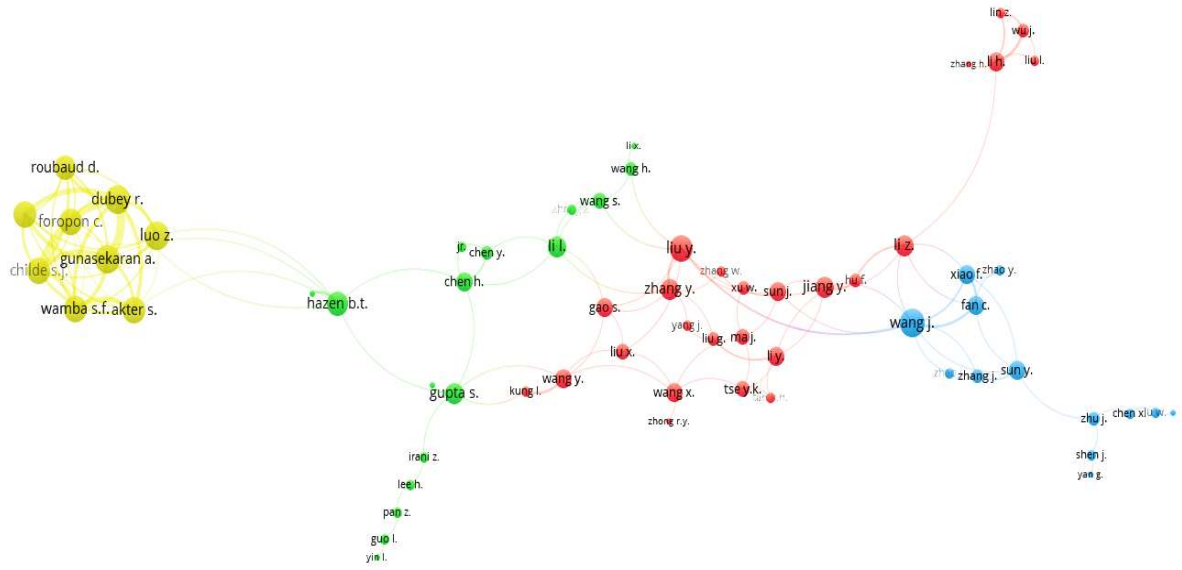


\* Minimum occurrence = 10

**Figure 4: Density diagram of index keywords\***



\* Minimum occurrence = 10



**Figure 5: Network of authors from co-authorship analysis\***

\* minimum publications = 3; minimum citations = 10

**Figure 6: Network of organisations from co-authorship analysis**

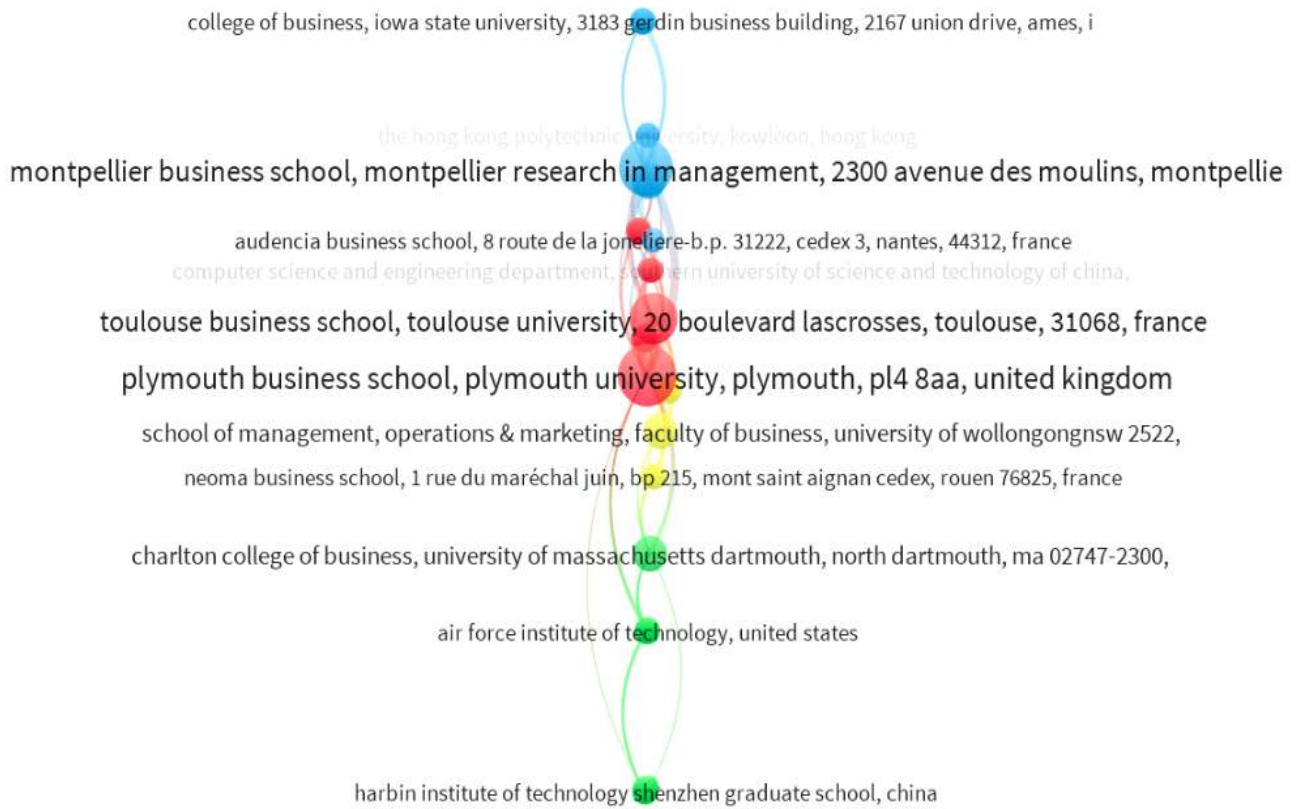
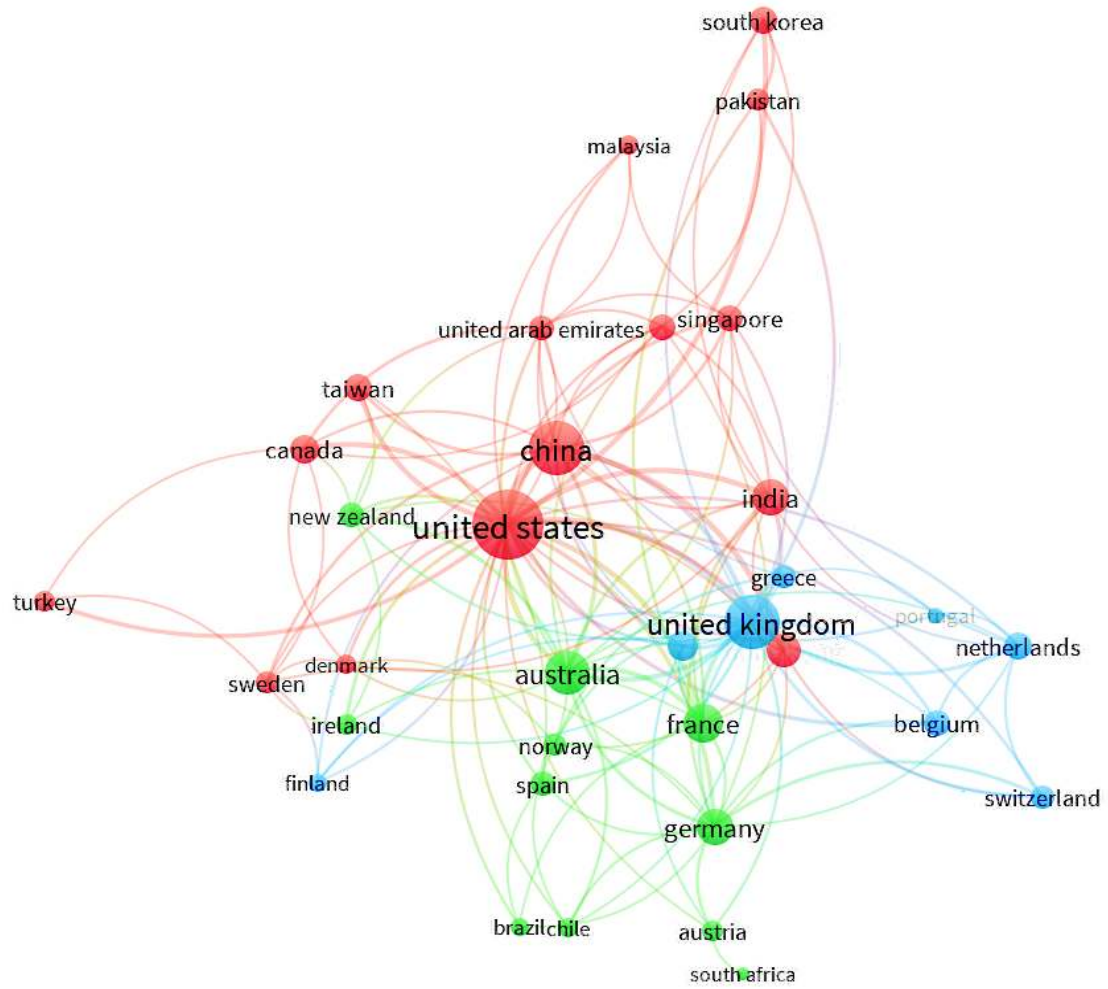


Figure 7: Network of countries from co-authorship analysis\*



\* minimum publications = 10; minimum citations = 100



**Table 16: Literature classification from co-citation analysis**

Cluster*	Thematic area	Journal	Articles
<b>Cluster 1</b> (Color code: Pink; N=40)	Inputs for strategic decision making	International Journal of Production Economics	7
		Journal of Business Research	4
		Journal of Business Logistics	3
<b>Cluster 2</b> (Color code: Green; N=35)	Concept development for big data analytics	Harvard Business Review	4
		Decision Support Systems	2
		Communications of the ACM	2
<b>Cluster 3</b> (Color code: Black; N=19)	Trends in applications of big data analytics	International Journal of Information Management	3
		European Journal of Information Systems	2
		International Journal of Medical Informatics	1
<b>Cluster 4</b> (Color code: Orange; N=16)	Efficient supply chain management	International Journal of Production Economics	5
		Journal of Business Research	2
		International Journal of Information Management	2

\* The clusters are numbered based on their size, not the sequence of their emergence

**Table 17: Top 10 articles in each cluster from prestige analysis**

Articles in cluster 1	PageRank score	Articles in cluster 2	PageRank score
Waller and Fawcett (2013)	0.012741	George, Haas, and Pentland (2014)	0.017412
Akter et al. (2016)	0.011086	McAfee and Brynjolfsson (2012)	0.014210
Erevelles, Fukawa, and Swayne (2016)	0.010952	Davenport (2013)	0.012844
Chae (2015)	0.009756	Trkman et al. (2010)	0.010887
Xu, Frankwick, and Ramirez (2016)	0.009489	Davenport and Patil (2012)	0.010142
Côrte-Real, Oliveira, and Ruivo (2017)	0.009176	Kitchin (2014)	0.010081
Wang, Kung, and Byrd (2018)	0.009003	Jagadish et al. (2014)	0.009726
Dubey et al. (2016)	0.008623	Davenport, Harris, and Shapiro (2010)	0.009535
Popovič et al. (2018)	0.008511	Acito and Khatri (2014)	0.009134
Sharma, Mithas, and Kankanhalli (2014)	0.008137	Siemens (2013)	0.008959

Articles in cluster 3	PageRank score	Articles in cluster 4	PageRank score
LaValle et al. (2011)	0.015055	Wamba et al. (2015)	0.018236
Gandomi and Haider (2015)	0.014237	Hazen et al. (2014)	0.016653
Chang, Kauffman, and Kwon (2014)	0.013637	Chen, Chiang, and Storey (2012)	0.016494
Lycett (2013)	0.011981	Wang et al. (2016)	0.015785
Chaudhuri, Dayal, and Narasayya (2011)	0.011604	Wamba et al. (2017)	0.014683
Raguseo (2018)	0.009209	Zhong et al. (2015).	0.013236
Yaqoob et al. (2016)	0.008705	Tan et al. (2015)	0.011908
Oztekin, Delen, and Kong (2009)	0.008276	Gunasekaran et al. (2017)	0.011818
Wang and Hajli (2017)	0.008195	Papadopoulos et al. (2017)	0.011522
Akter and Wamba (2016)	0.007861	Hu et al. (2014)	0.011242

**Figure 8: Evolution of clusters from dynamic co-citation analysis**



Figure 8.1: 2009 – 2010

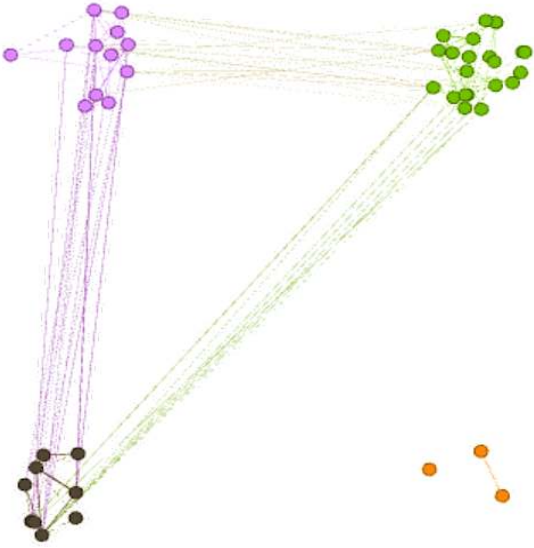


Figure 8.2: 2009 – 2013

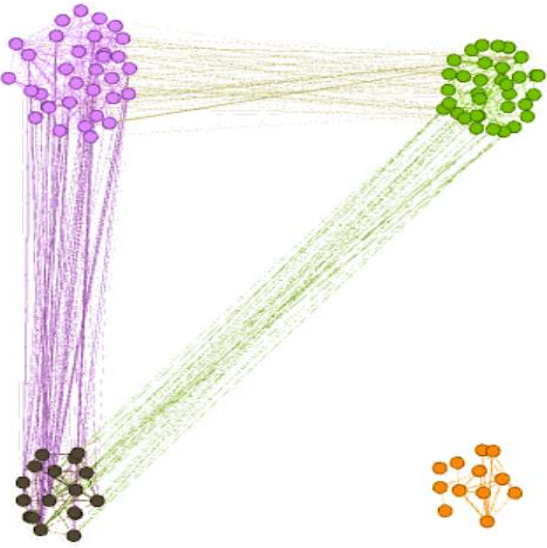


Figure 8.3: 2009 – 2016

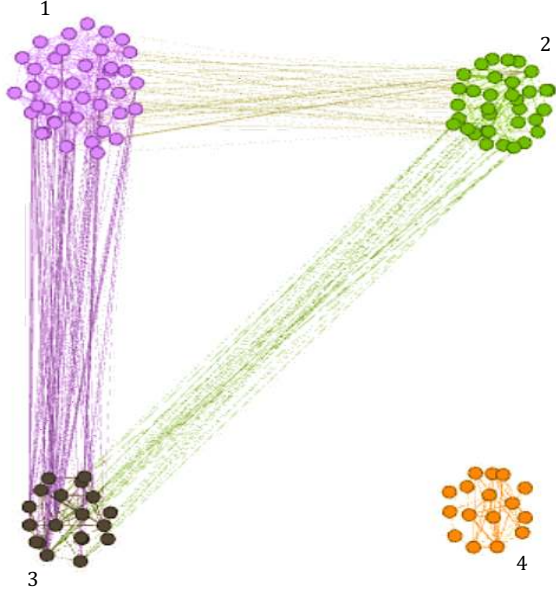
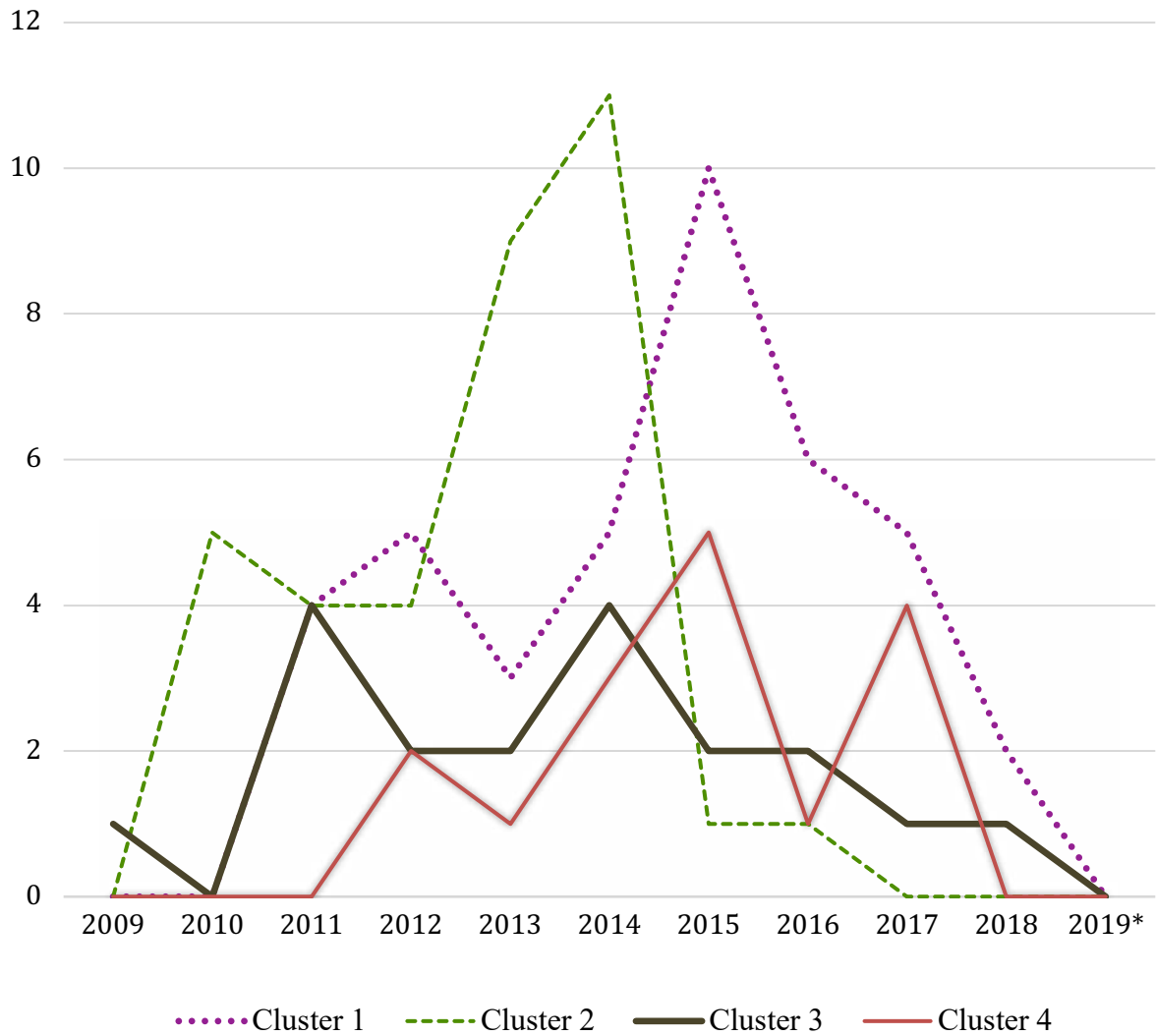


Figure 8.4: 2009 – 2019

Data source: Exhibit A

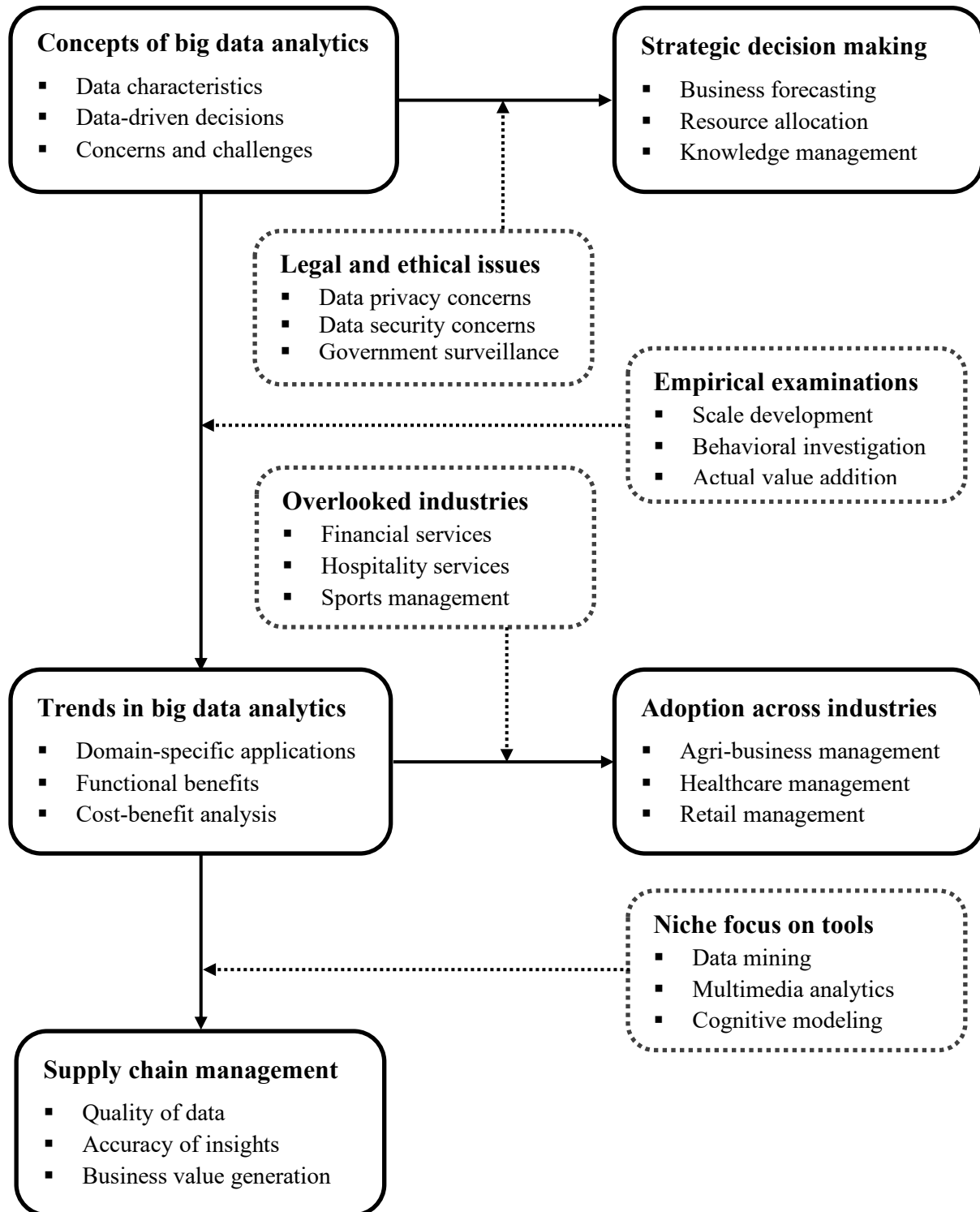
**Figure 9: Number of articles added to a cluster in a year**



\* as of August 01, 2019

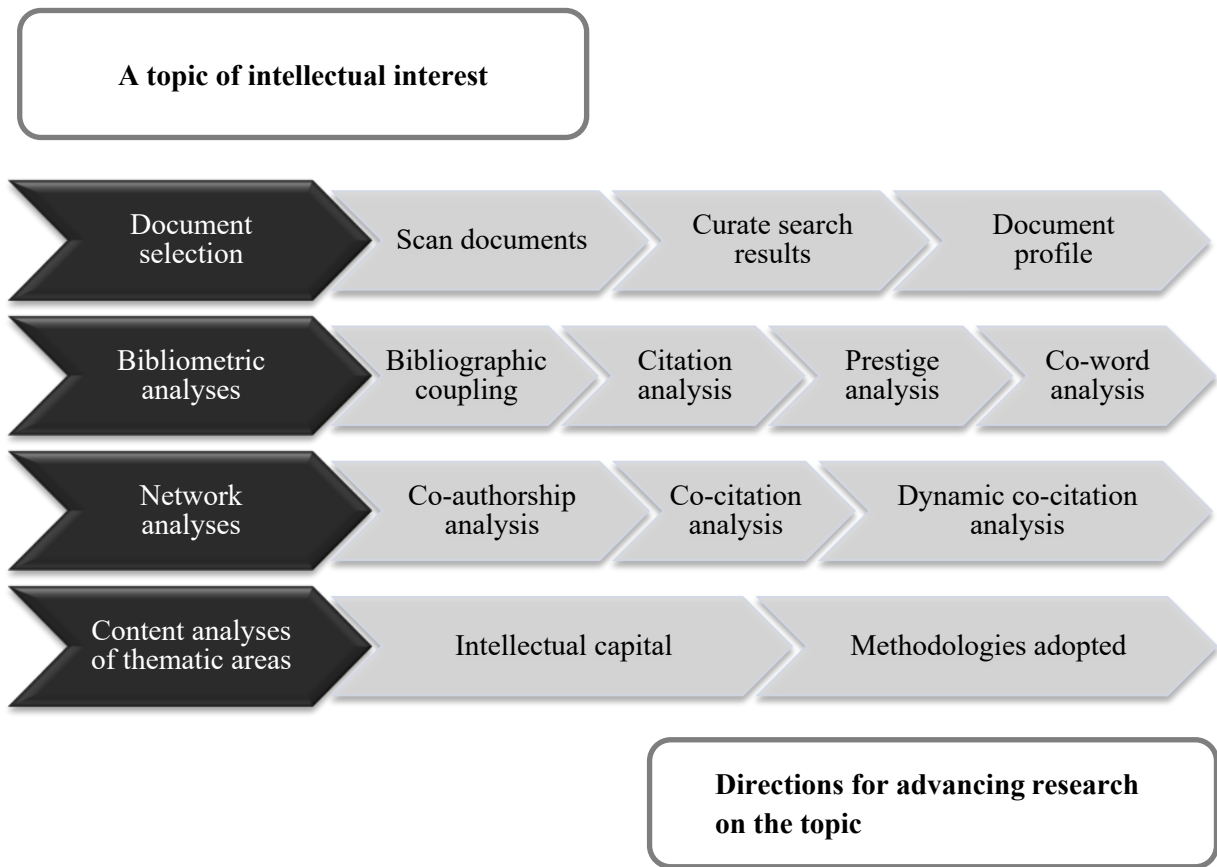
Data source: Exhibit B

**Figure 10: A conceptual framework for big data analytics in enterprises**



Note: Solid lines represent themes from prior research. Dotted lines represent future research scopes.

## Exhibit A: Protocol for a bibliometric study\*



[\* The protocol summarises methodologies followed in prior research (Caviggioli and Ughetto, 2019; Fahimnia, Sarkis, and Davarzani 2015; Xu et al., 2018)]

**Exhibit B: Evolution of clusters from dynamic co-citation analysis**

Year	Number of articles published			
	Cluster 1	Cluster 2	Cluster 3	Cluster 4
2009	0	0	1	0
2010	0	5	0	0
2011	4	4	4	0
2012	5	4	2	2
2013	3	9	2	1
2014	5	11	4	3
2015	10	1	2	5
2016	6	1	2	1
2017	5	0	1	4
2018	2	0	1	0
2019*	0	0	0	0
Total*	40	35	19	16

\* as of August 01, 2019