Kent Academic Repository

Technology in the 21st Century: New Challenges and Opportunities

ABSTRACT

Although big data, big data analytics (BDA) and business intelligence have attracted growing attention of both academics and practitioners, a lack of clarity persists about how BDA has been applied in business and management domains. In reflecting on Professor Ayre's contributions, we want to extend his ideas on technological change by incorporating the discourses around big data, BDA and business intelligence. With this in mind, we integrate the burgeoning but disjointed streams of research on big data, BDA and business intelligence to develop unified frameworks. Our review takes on both technical and managerial perspectives to explore the complex nature of big data, techniques in big data analytics and utilisation of big data in business and management community. The advanced analytics techniques appear pivotal in bridging big data and business intelligence. The study of advanced analytics techniques and their applications in big data analytics led to identification of promising avenues for future research.

KEY WORDS: business intelligence; big data; big data analytics; advanced techniques; decision-making

1. INTRODUCTION

In the last two decades, technological breakthroughs have ushered in a new era for businesses and governments (Amankwah-Amoah, 2017; Ayres & Williams, 2004; You et al., 2018). As Ayres and Williams (2004, p. 316) observed, "much of the world is connected via sophisticated networks that allow volumes of text, images, sound, and video to be exchanged in an instant". The field of business analytics has been one of the rapid growing and promising areas of business intelligence (Chen et al. 2012; Sheng et al., 2017). Business intelligence means the "concepts and methods to improve business decision making by using fact-based support systems" (Lim et al., 2013, p. 17). In this process, business analytics is concerned with methods and techniques for dealing with data to assess the past or real-time business performance, and it plays a critical role in the interconnected world (Davenport, 2006). Indeed, past studies have demonstrated that improvement in decision-making depends on progress in analytical tools (Ayres, 1984, 1989). This provides a solid basis for future business planning and decision-making by offering data users better knowledge and insights into business activities.

Technological progress is critical to the global economy (Ayres, 1988; Ayres & Williams, 2004), especially with new applications of information and communications technologies (ICTs) (Ayres & Williams, 2004). In the age of the Internet of Things, the exponential growth of technologies has brought greater complexity to business analytics (Beath et al., 2012). With significant scale escalation and scope expansion, the tremendous explosion of information has been called "big data", which is not only characterised by extremely large volume, but also by significance in its variety, velocity, veracity, variability and value (Gandomi & Haider, 2015; Katal et al., 2013; Jin et al., 2015).

Technological advancement and transition into a data-driven culture are important drivers for economic growth in the digital era; however, the prospect depends on the application of ICTs (Ayres & Williams, 2004). A growing body of literature supports the view that data-driven approaches, business intelligence and analytics largely depend on all kinds of data collection, information extraction and analytics techniques (Turban et al., 2008; Watson & Wixom, 2007). Although research on applications of big data analytics (BDA) and technological development have grown exponentially in the last ten years or so (Russom, 2011; Ayres & Williams, 2004), there remains inadequate clarity about what advanced techniques have been developed given the challenges inherent in big data's nature and how BDA has been applied in the business domain and discussed in scholarly work.

Against this backdrop, the main purpose of this paper is to extend Professor Ayres' idea of the significance of technology in stimulating economic growth by recognising the importance of harnessing big data in the 21st century. We do so by surveying the literature on big data and BDA techniques in management application and outline directions for future research. This study is closely related to those of Chen et al. (2012) and Chen et al. (2014). One of our key arguments is that harnessing new technology and data to make better and informed decision can contribute to the wider discussion on different mechanisms for technological changes and driving economic growth and development.

Following the idea of big data value chain in Chen et al. (2014) and taxonomy of emerging analytics areas in Chen et al. (2012), we take on both technical and managerial perspectives to explore current research in the management community. We develop a conceptual framework illustrating that BDA is the key that bridges the gap between big data and business intelligence. To that end, we identify the key technological advancements that address big data challenges and depict how BDA has been discussed and applied in management research. By incorporating and clarifying big data, BDA and business intelligence, this study helps identify research and business opportunities. Although some studies have charted the historical evolution of big data (see Phillips, 2017), we steer away from such analysis to offer a more robust review of the current state of knowledge.

The rest of this paper is organised as follows. The next two sections clarify the key terms, scope and road map for this review. We then describe the findings on advanced technological enablers in BDA workflow and on BDA methods adopted in management applications. The final section presents the discussion and conclusion, highlighting the research gap and setting an agenda for future research on applying BDA in management.

2. TECHNOLOGY AND BIG DATA IN THE 21ST CENTURY

The term "big data" refers to the extremely large amount of structured, semi-structured, and unstructured data continuously generated from diversified sources, which inundates business operations in real time and impacts decision-making through mining knowledge from massive data (Phillips, 2017; Wamba et al., 2017. It presents a great opportunity to enhance our capabilities to better understand the world (Amankwah-Amoah, 2015, 2016). In addition, challenges inherent to big data are emerging that require technological advancement to help capture value from big data.

The proliferation of big data has inspired practitioners and academics to take advantage of it with more effective analytics (Phillips, 2017; Wamba et al., 2017). A host of factors such as diversified data sources and types, faster data generation speed and an urgent need of efficient analytics in such a data-driven business environment has motivated the advance in techniques to achieve better analytics performance. Indeed, technological breakthroughs and innovation have led to methodological improvements to perform complex data analysis over the past few years (McAfee et al., 2012; Davenport et al., 2012).

Advanced analytics, also known as discovery analytics or exploratory analytics, is a collection of techniques used to "discover new business facts that no one in the enterprise knew before" (Russom, 2012, p. 2). Applying advanced techniques to analysing complicated data sets has led to BDA (LaValle et al., 2011), which is comprised of advanced analytics techniques, hardware and software, platforms and tools to perform big data management and analysis. BDA is a combination of massive data sets and advanced analytics that investigates the specifically detailed aspects of business activities and provides on-going positions of the business (Russom, 2011; Wang et al., 2018). It provides great support for making decisions and taking actions based on evidence (Wang et al., 2018).

3. CONDUCTING THE REVIEW

As illustrated in Paré et al. (2015), a scoping review "attempts to provide an initial indication of the potential size and nature of the available literature on a particular topic" to "examine the extent, range and nature of research activities, determine the value of undertaking a full systematic review, or identify research gaps in the extant literature" (p. 186). This study is a mapping review that aims to clarify BDA research trends and identify the advanced techniques applied in current business intelligence. For this purpose, we adopted the best practice advocated by past studies (e.g., Cropanzano, 2009; Short, 2009; Webster & Watson, 2002) and employed by review studies such as Short et al. (2009) and Ucbasaran, Shepherd, Lockett, & Lyon (2013). We also followed the four-step process of doing content analysis as introduced in Seuring and Gold (2012).

The surveyed literature was comprised of English-written peer-reviewed papers on big-data-related topics covering a 16-year period from 2000-2015. We set up the review from the turn of this century, when big-data-related concepts started to gain prominence, although big data as a management-related term first gathered steam around the year of 2007 (Halevi & Moed,

2012). Given that research has flourished in both top-tier and lower-ranked journals, we decided not to limit the scope of the review to the status of journals in line with past reviews. In addition to articles published in academic journals, several conference proceedings, unpublished studies and book chapters were included.

To identify studies that describe techniques and their applications in BDA, extensive research in library services and major databases (including Informs, Business Source Complete, Sage, Wiley, Springer, Emerald, ScienceDirect, JSTOR, Other publishers' online service) was carried out. Structured keywords such as "big data", "big data analytics", "advanced analytics", and relevant terms of technologies and techniques were used to search and identify related studies. Depending on the sufficiency of search outcomes, combinations of keywords were used to extend or narrow selection results. In addition, to ensure the identified studies were within the scope of this review, a quick content check was conducted by reading the abstracts and considering the appropriateness of the topic based on the definitions in the previous section.

The next stage of the process entails a detailed examination of the whole paper and classification of selected papers according to the techniques applied in the study. Within the classification scheme, all papers are tagged with labels like the year of publication, authors, research methodology, key areas, key features and key findings. In all, 280 studies were identified and analysed in this study. Around 83% of the sample was from academic journals in management domains, while the remaining were conference papers, books or industrial research reports. Among the journal articles, most were in fields such as information management, marketing and operations, while there are fewer organisational studies, sector studies and general management subjects. Although it shows an imbalance in the distribution of subjects in the collected papers, studies into big data show a growing trend over the sample period, especially after 2010.

To delineate the findings, we utilised Figure 1 as a road map that links big data, BDA and business intelligence as cores in a single framework. Figure 1 shows the features and linkages between the various constructs, i.e., big data, BDA and business intelligence. At the first stage of this review, we focused on the technical aspect of BDA. We argued that the nature of big data presents technical challenges to data management and data analytics, but by applying advanced technologies and techniques at each stage of the analytics workflow, big data are manageable and valuable information can be potentially extracted to inform business decisions. Then we moved on to examine the application of advanced techniques in BDA for achieving

business intelligence. Based on data types, we classified big data into structured data and unstructured data, which corresponds to the classification of emerging BDA areas described in Chen et al. (2012). The focus at this stage was to look into the application of analytics techniques in management research.

Insert Figure 1 about here

----

4. TECHNOLOGY IN BIG DATA ANALYTICS WORKFLOW

Due to the challenges of big data in terms of huge volume, high variety, high velocity, high variability, low veracity and high value, greater efficiency is a primary goal in handling such data sets. Ayres (1984) emphasises that better analytical tools with enhanced capabilities are the keys to improving analysis and thus decision-making. The review suggests that advanced analytics techniques can improve efficiency at each stage of the BDA workflow. Indeed, new techniques have been constantly proposed and discussed in fields such as computer science and engineering. These techniques can help achieve better data quality, adequate storage space, faster access and process speed, deeper analysis and more concise results presentation. In the following discussions, various techniques are introduced based on analytics workflow structure and around a theme of analytics efficiency.

4.1 Big Data Acquisition

The review revealed that sources generating data have become more diversified and massive data may come from enterprise business, networking, scientific experiments and others (see Hu, Wen, Chua, & Li, 2014). Data acquisition not only concerns extracting raw data from data sources, but includes collection, transmission and pre-processing of data (Hu et al., 2014; Chen et al., 2014; Tsai et al., 2016). Khan et al. (2014) also indicate that data collection, filtering and classification are the main tasks at this stage. Facing massive data, targeting useful data and cleaning the data are particularly vital. Through these steps, analytics can be performed on data with better quality by adopting advanced techniques in the acquisition.

To retrieve raw data generated from various sources, new techniques are developed. A few commonly used methods seen in literature include log files (e.g., Moreta & Telea, 2007;

Suneetha & Krishnamoorthi, 2009; Thelwall, 2001b; Nicholas & Huntington, 2003), sensors (e.g., Wang & Liu, 2011; Jonsson & Eklundh, 2002; Luo et al., 2009), web crawlers (e.g., Choudhary et al., 2012; Kaplan & Haenlein, 2010; Thelwall, 2001a), mobile devices (Baak et al., 2013; Kaplan & Hegarty, 2005; Laurila et al., 2012) and RFID (Roberts, 2006). The distributed architecture is normally used to capture system log, and web crawler is often used for unstructured network data (Liu et al., 2013).

Data transmission refers to transporting the data into a data storage infrastructure or data centre. It has IP backbone transmission and data centre transmission. The inter-DCN (e.g., Ghani et al., 2000; Shieh, 2011) and intra-DCN (e.g., Barroso et al., 2013) allow mobility of data within or between storage devices. A final step and a key to improving data quality is data pre-processing, which requires data integration (e.g., Lenzerini, 2002; Gravano et al., 2003), cleansing (e.g., Maletic & Marcus, 2000; Jeffery et al., 2006) and redundancy elimination (e.g., Sarawagi & Bhamidipaty, 2002; Tsai & Lin, 2012; Hussain et al., 2010). It helps aggregate data into a uniform format and improves data consistency. The processed data have benefits in BDA due to reduced costs and increased availability (Chaudhuri et al., 2011).

4.2 Big Data Storage

One key observation is that data volume is experiencing spectacular growth, which raises the standard for data storage in terms of space and ease to access. The review indicates that advanced data storage techniques can help maximise storage space using a distributed or a networked infrastructure and help save costs and time by conducting analysis within the database or memory (Lv et al., 2017). Data storage capabilities rely on hardware infrastructure, database and management and programming models (Hu et al., 2014). To store larger data sets, hardware infrastructure have advanced to networking architecture, such as direct attached storage (DAS), network attached storage (NAS) and storage area network (SAN) (e.g., Chen et al., 2014; Khan et al., 2014; Shiroishi et al., 2009; Gibson & van Meter, 2000; Reed et al., 2000; Barker & Massiglia, 2002; Telikepalli et al., 2004). These storage facilities connect to each other or link to a network, which gives easier access to other storage space. Moreover, distributed storage frameworks such as Google GFS and Hadoop HDFS have advantages in dealing with large data sets and steaming data with cheaper disk drivers (e.g., Ghemawat et al., 2003; Shvachko et al., 2010). Data are broken down into smaller scales and distributed in different servers, which enables scalability for processing.

As for the database to store and manage data, RDBMS is mainly for large structured data sets (Moniruzzaman & Hossain, 2013; Ramakrishnan & Gehrke, 2000).

For non-relational databases, there are three major types: key-value database, document database and column-oriented database. It is faster to retrieve data with key and value pairs and easier to compress data and operate parallel processing with data records in a sequence of columns (e.g., DeCandia et al., 2007; Chodorow, 2013; Cattell, 2011; McCreary & Kelly, 2013). In-memory (e.g., Watson, 2014; Hahn & Packowski, 2015) and in-database (e.g., Russom, 2012; Chaudhuri et al., 2011) techniques enable data processing and analysis within memory or the database instead of transporting data between the data centre and disks. In addition, public or private clouds infrastructure help facilitate data management. Resources in clouds can be allocated dynamically, but this raises security concerns (e.g., Talia, 2013; Bi & Cochran, 2014; Assunção et al., 2015).

4.3 Big Data Processing

Elgendy and Elragal (2014) identify four requirements of processing big data: fast data loading, fast query processing, highly efficient utilisation of storage space and strong adaptability to highly dynamic workload patterns. The review found that timely processing techniques can speed up data processing and enhance the efficiency of large-scale data analytics. Researchers have begun to examine a range of big data techniques such as generic processing model (Grolinger et al., 2014; Dean & Ghemawat, 2008, 2010; Condie et al., 2010; Sagiroglu & Sinanc, 2013), stream processing model (e.g., Neumeyer et al., 2010; Cherniack et al., 2003; Stonebraker et al., 2005), and graph processing model (Malewicz et al., 2010; Lumsdaine et al., 2007; Salihoglu & Widom, 2013). The latter two can deal with large-scale data using graph and event nodes.

In terms of querying, compared to relational processing that uses Structured Query Language (SQL) to access structured data in the relational database (e.g., Pedersen & Jensen, 2001), parallel processing can perform more efficient queries. Data are spread across many servers, and computing problems are solved on separate servers in parallel. No memory or resources need to be shared across different servers, and it is easy to expand with additional servers. It is highly efficient for large-scale data sets and unstructured data (e.g., Roosta, 2012; Jordan & Alaghband, 2002; Parhami, 2006).

4.4 Big Data Analysis

Data analysis seeks to "understand the relationships among features" and "develop effective methods of data mining that can accurately predict future observations" (Khan et al., 2014, p. 10). It can be descriptive, predictive and prescriptive (Hu et al., 2014). The review revealed that three broad categories of techniques have also been adopted in big data analysis. Mature analysis methods including regression and statistical analysis have been used in advanced analytics, and they are widely adopted in analysing large data clusters. Furthermore, new approaches are emerging, and growing adoption of the advanced methods has been seen in BDA to facilitate decisions.

In particular, data mining (see Wu et al., 2008; Han et al., 2011; Witten et al., 2011) and machine learning (Singh, 2014) are gaining popularity. Given the fact that big data are large in volume and diversified in format, data analysis relies more on computational algorithms to conduct in-depth examinations. Najafabadi et al. (2015) point out that deep learning algorithms benefit extracting information from massive data. Moreover, to present the massive amount of data and analysis results, more concise and interactive methods and platforms are required, such as advanced data visualisation (ADV). These platforms are relatively new in business intelligence and analytics, but they are helpful in give data users better engagement with data analysis and interpretation.

5. BIG DATA ANALYTICS TECHNIQUES IN MANAGEMENT APPLICATION

Having set out the various advanced techniques in data analytics workflow, this section relates to BDA mainly in management disciplines. Following the classification approach in Chen et al. (2012), we group the literature based on various data types: structured data, text, web and multimedia data, network data and mobile data. The review indicates that techniques for structured data analytics have seen tremendous improvements. Novel platforms and advanced analytics have been developed, while several mature programs in current business intelligence system continue serving in BDA by continuously optimising the designs for handling large volume data. The review also reveals that current research emphasises unstructured data analytics, and the majority of previous studies apply advanced techniques to analyse big data for enhancing management effectiveness.

In the business intelligence 1.0 system (Chen et al., 2012), data are mostly structured with pre-defined index or primary keys indicating relationships of data entities. In business intelligence 2.0 and 3.0 systems, data are generated with faster speed and greater diversity, which requires new analytics techniques to meet new challenges in structured and unstructured data analytics. Given the significant interest in unstructured data analytics in the current business intelligence and analytics research (Hu et al., 2014, Chen et al., 2014), our discussions underscore such importance by discovering major themes in management-related big data research.

5.1 Text Analytics

A considerable stream of research investigates the topic of text analytics (see a detailed summary of literature in Appendix A). In our literature sample, many studies target text mining, clustering and classification to detect topics and opinions, personalisation and recommendations and other areas. Text analytics depends largely on text mining, which deals with unstructured text from documents, emails, logs, web pages, social media, comments, feedback and so on. In this process, document representation and query processing help retrieve information, while Natural Language Processing (NLP) detects certain words, phrases, events and topics from massive data. Then the processed data can be used to construct models for further analysis, such as topic models (e.g., Blei, 2012), language models (e.g., Kao & Poteet, 2007), sentiment or opinion mining (e.g., Pang & Lee, 2008; Garg & Chatterjee, 2014; Ibrahim et al. 2017), interactive question and answer systems (e.g., Fan et al., 2006). Hence, information retrieval and statistical NLP are the bases from which more text-based searching and analysing techniques are innovated. These novel approaches can serve as techniques for other emerging research such as web and network analytics.

5.2 Web and Multimedia Analytics

Web analytics has become an essential part in Web 2.0 systems (O'Reilly, 2007). It aims to discover and analyse useful information from web documents and services such as web text, link structure, web logs and other types of web data (Ashton et al., 2014). The review found that web mining is a big part of web analytics (see Table 1). The web content, structure and usage mining data (Pal et al., 2002) provide abundant information about user dialogues and behaviours (e.g., clickstream), which is useful for improving business operations, particularly for improving the efficiency of marketing activities in electronic commerce.

Insert Table 1 about here

In addition to web data presented in text format, other information contained in images, audio and video is becoming new research objectives. Current research mainly concentrates on the multimedia summarisation (Ding et al., 2012), multimedia annotation (Wang et al., 2012), multimedia index and retrieval (Lew et al., 2006) and multimedia recommendation (Park & Chang, 2009). These techniques help extract hidden information from images, audios and videos with proper classification, annotation and retrieval, which can be used with other web and media data to suggest particular content to users based on their preference. However, we noted a handful of studies using multimedia data for gaining business insights, which lags behind other data exploration. It may due to the richness of content, which causes difficulties in extracting information and organising it into structured data.

5.3 Network Analytics

Network science has evolved with the rapid growth of online interactions and online social networking since 2000. Such massive amounts of user-generated data often reflect consumers' opinion and connections. The review illustrates that social media analysis seems to be a promising direction in big data research. Many studies use social media data to discover consumers' sentiment, predict consumers' behaviour, detect relationships and influences in the online community and enhance brand and sale performance (see Appendix B). Discovery of potential links, social influence and interactions in the network is helpful for understand the preferences, behaviours and dynamics in the virtual community.

5.4 Mobile Analytics

Mobile devices are penetrating universally, and mobile computing technologies enable millions of applications to generate a vast amount of information. Mobile and sensor-based systems foresee a promising future in business intelligence and analytics (see Table 1). Gathered from applications embedded in smart devices such as smartphones and tablets, data generated from these sources are usually fine-grained, location-specific, context-aware and highly personalised. The review shows that typically, mobile analytics uses mobile sensing, web logs, networking and applications to acquire personalised data from mobile users and

create behavioural models of individuals to realise customised advertisements or recommendations. It provides great opportunities to innovate and advance understanding of markets and customers in a timely manner.

6. DISCUSSIONS AND CONCLUDING THOUGHTS

Although the last ten years has witnessed a surging stream of research on big data, BDA and business intelligence, limited attention has been paid to harnessing big data to improve managerial and economic decisions. This study sought to fill this void by reviewing the literature on BDA across the social sciences and thus shed light on how data have been utilised in different settings. The paper identified 280 studies published in the past 16 years and classified them with a structured scheme. This review does not claim to be complete, given the fast growing body of research in relevant areas. In reflecting on Professor Ayre's contributions, we saw the need to extend his works by incorporating the richness of big data research. Clarifying the current knowledge of advanced analytics techniques benefits future application of BDA in the business sector. This study serves as the link between our survey of literature and Professor Ayres' emphasis on the role of technology in making better decisions. Within the review scope, an investigation into the collected studies led to the identification of two dimensions in current BDA research. From a technical perspective, studies focus on the analytics workflow, in which advanced techniques can help improve efficiency in handling big data. From a managerial perspective, advanced techniques target diversified data types and sources and are used in business analytics to achieve better managerial insight.

The review of the multi-disciplinary literature on BDA techniques and their applications further revealed several underexplored areas. One key observation in this review is that among the four topics in unstructured data analytics, text analytics and network analytics seem to be more attractive to management researchers than web, multimedia and mobile data techniques. This may be a result of the growing number of users on social media platforms (e.g., Facebook and Twitter) and accessibility of the data sources that sparks such big data research (Rapp et al., 2013; Mount & Martinez 2014; Chan et al., 2017). While text and network analytics are popular, analysis of web data, audio, video, mobile and sensor data are rarely seen in general management studies. Nevertheless, other types of unstructured data are also being generated and collected on an astronomical scale across various industry sectors, and most organisations have more data than they know how to use effectively (LaValle et al., 2011).

Perhaps the most prominent gap in the current literature is the limited attention paid by scholars in information science and management science to the potentials of harnessing BDA to achieve competitive advantage. One shortcoming we spotted is that the majority of studies have been conducted by scholars in areas such as marketing and operations management. Information management has yielded a large number of papers proposing novel approaches to handling big data, and the applications of these techniques are seen in marketing and operations research. Studies in this area cannot flourish in isolation and require research inputs from other management disciplines such as strategy, innovation, entrepreneurship, international business, organisation and sector studies. Therefore, research is needed to advance further understanding and utilisation of BDA in managerial applications. Nonetheless, in other management subjects, the influence of big data on the performance of management is understudied and offers promising avenues for future research. Similarly, there is an urgent need for closer collaborations between the academics and the industry to advance big data research and applications.

In addition, empirical and modelling papers are in the majority of studies reviewed, which tend to utilise advanced techniques and propose novel approaches to processing big data. The volume, variety and velocity characteristics of big data require a combination of interdisciplinary knowledge and a mixed methodological approach to address the challenges of using it (Chan et al., 2016). Despite the fact that big data research has seen rapid growth in management disciplines, it is still at an early stage and many key questions remained unanswered.

Big data analytics can empower businesses to perform better predictive and prescriptive analysis to forecast and plan for the future, which are essential for management to make accurate decisions (see also Ayres, 1989). To guide future research, a framework is proposed to identify opportunities (see Figure 2). In this cycle, big data, BDA and business intelligence are linked, and research opportunities can be explored from two alternative routes. One is the top-down strategy. Guided by the business objective and intelligence desired, proper analytics techniques can be selected based on big data theory and applied to available data sets, which will help make better use of big data. On the other hand, a bottom-up strategy indicates that research can start with sorting out available big data and then analysing it with suitable techniques and technology support to gain valuable insights, thereby ultimately making an impact on business decisions and performance.

Insert Figure 2 about here

One limitation of this study is the selected scope using a keyword search approach. It may omit some earlier studies that might not use certain keywords that were not popular at that time. Future research opportunities can be identified in each element and linkage in this framework, which can help identify new research interests and may also offer implications for enterprises to take advantages of big data.

Insert Table 2 about here

Table 2 proposes several unanswered questions in present BDA research in management field. Among them, we think the priority should be developing a clearer picture of big data potentials and BDA methods. Given that there is a lack of theories and conceptual maps to instruct researchers around big data themes, guidance on how to leverage big data and analytics techniques for achieving business intelligence are in urgent need to systematise the research with solid theoretical foundations. In honour of Professor Robert Ayres' contributions to research in technological forecasting and social change, we hope that this study can serve as a useful reference point for researchers in management fields to advance big data research, forecasting and applications of BDA.

REFERENCES

Abrahams, A., Fan, W., Wang, G., Zhang, Z., & Jiao, J. (2014). An integrated text analytic framework for product defect discovery. *Production and Operations Management*, *24*(6), 975-990.

Alfaro, C., Cano-Montero, J., Gómez, J., Moguerza, J., & Ortega, F. (2013). A multi-stage method for content classification and opinion mining on weblog comments. *Annals of Operations Research*, *236*(1), 197-213.

Aliguliyev, R. (2009a). A new sentence similarity measure and sentence based extractive technique for automatic text summarization. *Expert Systems with Applications*, *36*(4), 7764-7772.

Aliguliyev, R. (2009b). Clustering of document collection – A weighting approach. *Expert Systems with Applications*, *36*(4), 7904-7916.

Amankwah-Amoah, J. (2015). Safety or no safety in numbers? Governments, big data and public policy formulation. *Industrial Management and Data System*, *115*(9), 1596–1603.

Amankwah-Amoah, J. (2016). Emerging economies, emerging challenges: Mobilising and capturing value from big data. *Technological Forecasting and Social Change*, *10*, 167-174.

Amankwah-Amoah, J. (2017). Integrated vs. add-on: A multidimensional conceptualisation of technology obsolescence. *Technological Forecasting and Social Change*, *116*, 299–307.

Andrews, M., Luo, X., Fang, Z., & Ghose, A. (2016). Mobile ad effectiveness: Hyper-contextual targeting with crowdedness. *Marketing Science*, *35*(2), 218-233.

Archak, N., Ghose, A., & Ipeirotis, P. (2011). Deriving the pricing power of product features by mining consumer reviews. *Management Science*, *57*(8), 1485-1509.

Ashton, T., Evangelopoulos, N., & Prybutok, V. (2013). Extending monitoring methods to textual data: a research agenda. *Quality & Quantity*, *48*(4), 2277-2294.

Assunção, M., Calheiros, R., Bianchi, S., Netto, M., & Buyya, R. (2015). Big data computing and clouds: Trends and future directions. *Journal of Parallel and Distributed Computing*, *79-80*, 3-15.

Autio, E., Dahlander, L., & Frederiksen, L. (2013). Information exposure, opportunity evaluation, and entrepreneurial action: An investigation of an online user community. *Academy of Management Journal*, *56*(5), 1348-1371.

Ayres, R. U. (1984). Improving the scientific basis of public and private decision-making. *Technological Forecasting and Social Change, 26*(2), 195-199.

Ayres, R. U. (1988). Technology: The wealth of nations. *Technological Forecasting and Social Change, 33*(3), 189-201.

Ayres, R. U. (1989). The future of technological forecasting. *Technological Forecasting and Social Change, 36*(1-2), 49-60.

Ayres, R. U., & Williams, E. (2004). The digital economy: Where do we stand?. *Technological Forecasting and Social Change, 71*(4), 315-339.

Baak, A., Müller, M., Bharaj, G., Seidel, H. P., & Theobalt, C. (2013). A data-driven approach for real-time full body pose reconstruction from a depth camera. In *Consumer Depth Cameras for Computer Vision* (pp. 71-98). Springer London.

Baek, H., Ahn, J., & Choi, Y. (2012). Helpfulness of online consumer reviews: Readers' objectives and review cues. *International Journal of Electronic Commerce*, *17*(2), 99-126.

Balahur, A., Hermida, J., & Montoyo, A. (2012). Detecting implicit expressions of emotion in text: A comparative analysis. *Decision Support Systems*, *53*(4), 742-753.

Balakrishnan, R., Qiu, X., & Srinivasan, P. (2010). On the predictive ability of narrative disclosures in annual reports. *European Journal of Operational Research*, *202*(3), 789-801.

Bao, Y. & Datta, A. (2014). Simultaneously discovering and quantifying risk types from textual risk disclosures. *Management Science*, *60*(6), 1371-1391.

Baralis, E., Cagliero, L., Jabeen, S., Fiori, A., & Shah, S. (2013). Multi-document summarization based on the Yago ontology. *Expert Systems with Applications*, *40*(17), 6976-6984.

Barker, R. & Massiglia, P. (2002). *Storage area network essentials*. New York: Wiley.

Barroso, L., Clidaras, J., & Hölzle, U. (2013). The datacenter as a computer: An introduction to the design of warehouse-scale machines, Second edition. *Synthesis Lectures on Computer Architecture*, *8*(3), 1-154.

Beath, C., Becerra-Fernandez, I., Ross, J., & Short, J. (2012). Finding value in the information explosion. *MIT Sloan Management Review*, *53*(4), 18-20.

Beebe, N., Clark, J., Dietrich, G., Ko, M., & Ko, D. (2011). Post-retrieval search hit clustering to improve information retrieval effectiveness: Two digital forensics case studies. *Decision Support Systems*, *51*(4), 732-744.

Berger, J. (2014). Word of mouth and interpersonal communication: A review and directions for future research. *Journal of Consumer Psychology*, *24*(4), 586-607.

Beverungen, A., Bohm, S., & Land, C. (2015). Free labour, social media, management: Challenging marxist organization studies. *Organization Studies*, *36*(4), 473-489.

Bi, Z. & Cochran, D. (2014). Big data analytics with applications. *Journal of Management Analytics*, *1*(4), 249-265.

Blei, D. (2012). Probabilistic topic models. *Communications of the ACM*, *55*(4), 77-84.

Cao, Q., Duan, W., & Gan, Q. (2011). Exploring determinants of voting for the "helpfulness" of online user reviews: A text mining approach. *Decision Support Systems*, *50*(2), 511-521.

Cascio, C., O'Donnell, M., Bayer, J., Tinney, F., & Falk, E. (2015). Neural correlates of susceptibility to group opinions in online word-of-mouth recommendations. *Journal of Marketing Research*, *52*(4), 559-575.

Castellanos, M., Gupta, C., Wang, S., Dayal, U., & Durazo, M. (2012). A platform for situational awareness in operational BI. *Decision Support Systems*, *52*(4), 869-883.

Cattell, R. (2011). Scalable SQL and NoSQL data stores. *ACM SIGMOD Record*, 39(4), 12-27.

Chan, H., Lacka, E., Yee, R., & Lim, M. (2017). The role of social media data in operations and production management. *International Journal of Production Research*, *55*(17), 5027-5036.

Chan, H., Wang, X., Lacka, E., & Zhang, M. (2016). A mixed-method approach to extracting the value of social media data. *Production and Operations Management*, *25*(3), 568-583.

Chatterjee, P., Hoffman, D., & Novak, T. (2003). Modeling the clickstream: Implications for web-based advertising efforts. *Marketing Science*, *22*(4), 520-541.

Chau, M. & Chen, H. (2008). A machine learning approach to web page filtering using content and structure analysis. *Decision Support Systems*, *44*(2), 482-494.

Chau, M. & Xu, J. (2007). Mining communities and their relationships in blogs: A study of online hate groups. *International Journal of Human-Computer Studies*, *65*(1), 57-70.

Chaudhuri, S., Dayal, U., & Narasayya, V. (2011). An overview of business intelligence technology. *Communications of the ACM*, *54*(8), 88-98.

Chen, C. & Tseng, Y. (2011). Quality evaluation of product reviews using an information quality framework. *Decision Support Systems*, *50*(4), 755-768.

Chen, H., Chiang, R. H., & Storey, V. C. (2012). Business intelligence and analytics: From big data to big impact. *MIS quarterly*, *36*(4), 1165-1188.

Chen, M., Mao, S., & Liu, Y. (2014). Big data: a survey. *Mobile Networks and Applications*, *19*(2), 171-209.

Cheng, Y. & Ho, H. (2015). Social influence's impact on reader perceptions of online reviews. *Journal of Business Research*, *68*(4), 883-887.

Cherniack, M., Balakrishnan, H., Balazinska, M., Carney, D., Cetintemel, U., Xing, Y., & Zdonik, S. B. (2003). Scalable distributed stream processing. In *CIDR* (Vol. 3, pp. 257-268).

Chodorow, K. (2013). *MongoDB: the definitive guide*. O'Reilly Media, Inc.

Chou, C. H., Sinha, A. P., & Zhao, H. (2010). A hybrid attribute selection approach for text classification. *Journal of the Association for Information Systems*, *11*(9), 491-518.

Choudhary, S., Dincturk, M. E., Mirtaheri, S. M., Moosavi, A., Von Bochmann, G., Jourdan, G. V., & Onut, I. V. (2012). Crawling rich internet applications: the state of the art. In *Proceedings of the 2012 Conference of the Center for Advanced Studies on Collaborative Research* (pp. 146-160). IBM Corp.

Chung, T., Wedel, M., & Rust, R. (2015). Adaptive personalization using social networks. *Journal of the Academy of Marketing Science*, *44*(1), 66-87.

Chung, W. & Tseng, T. (2012). Discovering business intelligence from online product reviews: A rule-induction framework. *Expert Systems with Applications*, *39*(15), 11870-11879.

Chung, W., Chen, H., & Nunamaker Jr, J. F. (2005). A visual framework for knowledge discovery on the Web: An empirical study of business intelligence exploration. *Journal of Management Information Systems*, *21*(4), 57-84.

Claussen, J., Kretschmer, T., & Mayrhofer, P. (2013). The Effects of Rewarding User Engagement: The Case of Facebook Apps. *Information Systems Research*, *24*(1), 186-200.

Colace, F., Casaburi, L., De Santo, M., & Greco, L. (2015). Sentiment detection in social networks and in collaborative learning environments. *Computers in Human Behavior*, *51*, 1061-1067.

Colace, F., De Santo, M., Greco, L., & Napoletano, P. (2014). Text classification using a few labeled examples. *Computers in Human Behavior*, *30*, 689-697.

Colace, F., De Santo, M., Greco, L., Moscato, V., & Picariello, A. (2015). A collaborative user-centered framework for recommending items in Online Social Networks. *Computers in Human Behavior*, *51*, 694-704.

Condie, T., Conway, N., Alvaro, P., Hellerstein, J. M., Elmeleegy, K., & Sears, R. (2010). MapReduce online. In *Nsdi* (Vol. 10, No. 4, pp. 20-33).

Costa, E., Ferreira, R., Brito, P., Bittencourt, I., Holanda, O., Machado, A., & Marinho, T. (2012). A framework for building web mining applications in the world of blogs: A case study in product sentiment analysis. *Expert Systems with Applications*, *39*(5), 4813-4834.

Coussement, K. & Van den Poel, D. (2008). Integrating the voice of customers through call center emails into a decision support system for churn prediction. *Information & Management*, *45*(3), 164-174.

Crawford Camiciottoli, B., Ranfagni, S., & Guercini, S. (2014). Exploring brand associations: an innovative methodological approach. *European Journal of Marketing*, *48*(5/6), 1092-1112.

Cropanzano, R. (2009). Writing Nonempirical Articles for Journal of Management: General Thoughts and Suggestions. *Journal of Management*, *35*(6), 1304-1311.

D'Haen, J., Van den Poel, D., & Thorleuchter, D. (2013). Predicting customer profitability during acquisition: Finding the optimal combination of data source and data mining technique. *Expert Systems with Applications*, *40*(6), 2007-2012.

Da Silva, N., Hruschka, E., & Hruschka, E. (2014). Tweet sentiment analysis with classifier ensembles. *Decision Support Systems*, *66*, 170-179.

Das, S. & Chen, M. (2007). Yahoo! for Amazon: Sentiment extraction from small talk on the web. *Management Science*, *53*(9), 1375-1388.

Davenport, T. H. (2006). Competing on analytics. *Harvard Business Review*, *84*(1), 98-107.

Davenport, T. H., Barth, P., & Bean, R. (2012). How big data is different. *MIT Sloan Management Review*, *54*(1), 43-46.

Dean, J. & Ghemawat, S. (2008). MapReduce. *Communications of the ACM*, *51*(1), 107-113.

Dean, J. & Ghemawat, S. (2010). MapReduce. *Communications of the ACM*, *53*(1), 72-77.

DeCandia, G., Hastorun, D., Jampani, M., Kakulapati, G., Lakshman, A., Pilchin, A., Sivasubramanian, S., Vosshall, P.W., & Vogels, W. (2007). Dynamo: amazon's highly available key-value store. *ACM SIGOPS Operating Systems Review*, *41*(6), 205-220.

Dehkharghani, R., Mercan, H., Javeed, A., & Saygin, Y. (2014). Sentimental causal rule discovery from Twitter. *Expert Systems with Applications*, *41*(10), 4950-4958.

Delen, D. & Crossland, M. (2008). Seeding the survey and analysis of research literature with text mining. *Expert Systems with Applications*, *34*(3), 1707-1720.

Delen, D. & Demirkan, H. (2013). Data, information and analytics as services. *Decision Support Systems*, *55*(1), 359-363.

Deng, Z., Luo, K., & Yu, H. (2014). A study of supervised term weighting scheme for sentiment analysis. *Expert Systems with Applications*, *41*(7), 3506-3513.

Ding, A., Li, S., & Chatterjee, P. (2015). Learning User Real-Time Intent for Optimal Dynamic Web Page Transformation. *Information Systems Research*, *26*(2), 339-359.

Ding, D., Metze, F., Rawat, S., Schulam, P. F., Burger, S., Younessian, E., Bao, L., Christel, M.G., & Hauptmann, A. (2012). Beyond audio and video retrieval: towards multimedia summarization. In *Proceedings of the 2nd ACM International Conference on Multimedia Retrieval* (Article No. 2). ACM.

Duan, J., Zhang, M., Jingzhong, W., & Xu, Y. (2011). A hybrid framework to extract bilingual multiword expression from free text. *Expert Systems with Applications*, *38*(1), 314-320.

Eisingerich, A., Chun, H., Liu, Y., Jia, H., & Bell, S. (2015). Why recommend a brand face-to-face but not on Facebook? How word-of-mouth on online social sites differs from traditional word-of-mouth. *Journal of Consumer Psychology*, *25*(1), 120-128.

Elgendy, N., & Elragal, A. (2014). Big data analytics: A literature review paper. In *Industrial Conference on Data Mining* (pp. 214-227). Springer International Publishing.

Fan, W., Gordon, M., & Pathak, P. (2006). An integrated two-stage model for intelligent information routing. *Decision Support Systems*, *42*(1), 362-374.

Fan, W., Wallace, L., Rich, S., & Zhang, Z. (2006). Tapping the power of text mining. *Communications of the ACM*, *49*(9), 76-82.

Fang, F., Dutta, K., & Datta, A. (2014). Domain adaptation for sentiment classification in light of multiple sources. *INFORMS Journal on Computing*, *26*(3), 586-598.

Fang, X., Hu, P., Li, Z., & Tsai, W. (2013). Predicting adoption probabilities in social networks. *Information Systems Research*, *24*(1), 128-145.

Feng, H., Tian, J., Wang, H., & Li, M. (2015). Personalized recommendations based on time-weighted overlapping community detection. *Information & Management*, *52*(7), 789-800.

Fersini, E., Messina, E., & Pozzi, F. (2014). Sentiment analysis: Bayesian ensemble learning. *Decision Support Systems*, *68*, 26-38.

Feuls, M., Fieseler, C., & Suphan, A. (2014). A social net? Internet and social media use during unemployment. *Work, Employment & Society*, *28*(4), 551-570.

Fong, N., Fang, Z., & Luo, X. (2015). Geo-conquesting: Competitive locational targeting of mobile promotions. *Journal of Marketing Research*, *52*(5), 726-735.

Fuller, C. M., Biros, D. P., & Delen, D. (2011). An investigation of data and text mining methods for real world deception detection. *Expert Systems with Applications, 38*(7), 8392-8398.

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

Gao, K., Xu, H., & Wang, J. (2015). A rule-based approach to emotion cause detection for Chinese micro-blogs. *Expert Systems with Applications*, *42*(9), 4517-4528.

García-Cumbreras, M., Montejo-Ráez, A., & Díaz-Galiano, M. (2013). Pessimists and optimists: Improving collaborative filtering through sentiment analysis. *Expert Systems with Applications*, *40*(17), 6758-6765.

García-Moya, L., Kudama, S., Aramburu, M., & Berlanga, R. (2013). Storing and analysing voice of the market data in the corporate data warehouse. *Information Systems Frontiers*, *15*(3), 331-349.

Garg, R., Smith, M., & Telang, R. (2011). Measuring information diffusion in an online community. *Journal of Management Information Systems*, *28*(2), 11-38.

Garg, Y., & Chatterjee, N. (2014). Sentiment analysis of twitter feeds. In *International Conference on Big Data Analytics* (pp. 33-52). Springer International Publishing.

Ghani, N., Dixit, S., & Wang, T. S. (2000). On IP-over-WDM integration. *IEEE Communications Magazine*, *38*(3), 72-84.

Ghemawat, S., Gobioff, H., & Leung, S. (2003). The Google file system. *ACM SIGOPS Operating Systems Review*, *37*(5), 29-43.

Ghose, A. & Han, S. (2011). An empirical analysis of user content generation and usage behavior on the mobile internet. *Management Science*, *57*(9), 1671-1691.

Ghose, A. & Han, S. (2014). Estimating demand for mobile applications in the new economy. *Management Science*, *60*(6), 1470-1488.

Ghose, A., Goldfarb, A., & Han, S. (2012). How is the mobile internet different? Search costs and local activities. *Information Systems Research*, *24*(3), 613-631.

Ghose, A., Ipeirotis, P., & Li, B. (2012). Designing ranking systems for hotels on travel search engines by mining user-generated and crowdsourced content. *Marketing Science*, *31*(3), 493-520.

Gibson, G. & Van Meter, R. (2000). Network attached storage architecture. *Communications of the ACM*, *43*(11), 37-45.

Glancy, F. & Yadav, S. (2011). A computational model for financial reporting fraud detection. *Decision Support Systems*, *50*(3), 595-601.

Goes, P., Lin, M., & Au Yeung, C. (2014). "Popularity effect" in user-generated content: Evidence from online product reviews. *Information Systems Research*, *25*(2), 222-238.

Goh, K., Heng, C., & Lin, Z. (2013). Social media brand community and consumer behavior: Quantifying the relative impact of user- and marketer-generated content. *Information Systems Research*, *24*(1), 88-107.

Gopaldas, A. (2014). Marketplace Sentiments. *Journal of Consumer Research*, *41*(4), 995-1014.

Gopinath, S., Chintagunta, P., & Venkataraman, S. (2013). Blogs, advertising, and local-market movie box office performance. *Management Science*, *59*(12), 2635-2654.

Gravano, L., Ipeirotis, P. G., Koudas, N., & Srivastava, D. (2003). Text joins in an RDBMS for web data integration. In *Proceedings of the 12th international conference on World Wide Web* (pp. 90-101). ACM.

Grolinger, K., Hayes, M., Higashino, W. A., L'Heureux, A., Allison, D. S., & Capretz, M. A. (2014). Challenges for mapreduce in big data. In *2014 IEEE World Congress on Services* (pp. 182-189). IEEE.

Guerreiro, J., Rita, P., & Trigueiros, D. (2016). A text mining-based review of cause-related marketing Literature. *Journal of Business Ethics*. *139*(1), 111-128.

Guo, Z., Wong, W., & Guo, C. (2014). A cloud-based intelligent decision-making system for order tracking and allocation in apparel manufacturing. *International Journal of Production Research*, *52*(4), 1100-1115.

Haenlein, M. (2011). A social network analysis of customer-level revenue distribution. *Marketing Letters*, *22*(1), 15-29.

Hahn, G. & Packowski, J. (2015). A perspective on applications of in-memory analytics in supply chain management. *Decision Support Systems*, *76*, 45-52.

Halevi, G., & Moed, H. (2012). The evolution of big data as a research and scientific topic: Overview of the literature. *Research Trends*, *30*(1), 3-6.

Han, J., Kamber, M., and Pei, J. (2011). *Data mining: Concepts and techniques.* Elsevier.

Hashimi, H., Hafez, A., & Mathkour, H. (2015). Selection criteria for text mining approaches. *Computers in Human Behavior*, *51*, 729-733.

He, W. (2013a). Examining students' online interaction in a live video streaming environment using data mining and text mining. *Computers in Human Behavior*, *29*(1), 90-102.

He, W. (2013b). Improving user experience with case-based reasoning systems using text mining and Web 2.0. *Expert Systems with Applications*, *40*(2), 500-507.

He, W., Wu, H., Yan, G., Akula, V., & Shen, J. (2015). A novel social media competitive analytics framework with sentiment benchmarks. *Information & Management*, *52*(7), 801-812.

Hennig-Thurau, T., Wiertz, C., & Feldhaus, F. (2014). Does Twitter matter? The impact of microblogging word of mouth on consumers' adoption of new movies. *Journal of the Academy of Marking Science*, *43*(3), 375-394.

Hildebrand, C., Häubl, G., Herrmann, A., & Landwehr, J. (2013). When social media can be bad for you: Community feedback stifles consumer creativity and reduces satisfaction with self-designed products. *Information Systems Research*, *24*(1), 14-29.

Ho, S., Bodoff, D., & Tam, K. (2011). Timing of adaptive web personalization and its effects on online consumer behavior. *Information Systems Research*, *22*(3), 660-679.

Homburg, C., Ehm, L., & Artz, M. (2015). Measuring and managing consumer sentiment in an online community environment. *Journal of Marketing Research*, *52*(5), 629-641.

Hu, H., Wen, Y., Chua, T. S., & Li, X. (2014). Toward scalable systems for big data analytics: A technology tutorial. *IEEE Access, 2*, 652-687.

Hu, N., Bose, I., Koh, N., & Liu, L. (2012). Manipulation of online reviews: An analysis of ratings, readability, and sentiments. *Decision Support Systems*, *52*(3), 674-684.

Huang, T. & Van Mieghem, J. (2013). Clickstream data and inventory management: Model and empirical analysis. *Production and Operations Management*, *23*(3), 333-347.

Hussain, T., Asghar, S., & Masood, N. (2010). Web usage mining: A survey on preprocessing of web log file. In *Information and Emerging Technologies (ICIET), 2010 International Conference on* (pp. 1-6). IEEE.

Hyung, Z., Lee, K., & Lee, K. (2014). Music recommendation using text analysis on song requests to radio stations. *Expert Systems with Applications*, *41*(5), 2608-2618.

Ibrahim, N.F., Wang, X. and Bourne, H., (2017). Exploring the effect of user engagement in online brand communities: Evidence from Twitter. *Computers in Human Behavior*, 72, 321-338.

Iyer, G. & Katona, Z. (2016). Competing for attention in social communication markets. *Management Science*, *62*(8), 2304-2320.

Janasik, N., Honkela, T., & Bruun, H. (2009). Text mining in qualitative research: Application of an unsupervised learning method. *Organizational Research Methods*, *12*(3), 436-460.

Jang, H., Sim, J., Lee, Y., & Kwon, O. (2013). Deep sentiment analysis: Mining the causality between personality-value-attitude for analyzing business ads in social media. *Expert Systems with Applications*, *40*(18), 7492-7503.

Järvinen, J. & Karjaluoto, H. (2015). The use of Web analytics for digital marketing performance measurement. *Industrial Marketing Management*, *50*, 117-127.

Jeffery, S. R., Alonso, G., Franklin, M. J., Hong, W., & Widom, J. (2006). A pipelined framework for online cleaning of sensor data streams. In *Data Engineering, Proceedings of the 22nd International Conference on* (pp. 140-140). IEEE.

Jin, X., Wah, B., Cheng, X., & Wang, Y. (2015). Significance and challenges of big data research. *Big Data Research*, *2*(2), 59-64.

Johnson, S., Safadi, H., & Faraj, S. (2015). The emergence of online community leadership. *Information Systems Research*, *26*(1), 165-187.

Jonsson, P. & Eklundh, L. (2002). Seasonality extraction by function fitting to time-series of satellite sensor data. *IEEE Transactions on Geoscience and Remote Sensing*, *40*(8), 1824-1832.

Jordan, H. & Alaghband, G. (2002). *Fundamentals of parallel processing*. Upper Saddle River, NJ: Prentice Hall/Pearson Education.

Kang, D. & Park, Y. (2014). Review-based measurement of customer satisfaction in mobile service: Sentiment analysis and VIKOR approach. *Expert Systems with Applications*, *41*(4), 1041-1050.

Kao, A. & Poteet, S. (2007). *Natural language processing and text mining*. New York: Springer.

Kaplan, A. & Haenlein, M. (2010). Users of the world, unite! The challenges and opportunities of Social Media. *Business Horizons*, *53*(1), 59-68.

Kaplan, E., and Hegarty, C. (2005). *Understanding GPS: Principles and Applications*. MA: Artech house.

Katal, A., Wazid, M., & Goudar, R. H. (2013). Big data: issues, challenges, tools and good practices. In *Contemporary Computing (IC3), 2013 Sixth International Conference on* (p. 404-409). IEEE.

Kayser, V., & Blind, K. (2016). Extending the knowledge base of foresight: The contribution of text mining. *Technological Forecasting and Social Change, 116,* 208-215.

Khan, F., Bashir, S., & Qamar, U. (2014). TOM: Twitter opinion mining framework using hybrid classification scheme. *Decision Support Systems*, *57*, 245-257.

Khan, N., Yaqoob, I., Hashem, I. A. T., Inayat, Z., Mahmoud Ali, W. K., Alam, M., Shiraz, M., & Gani, A. (2014). Big data: survey, technologies, opportunities, and challenges. *The Scientific World Journal*, *2014,* 1-18.

Kim, T., Hong, J., & Kang, P. (2015). Box office forecasting using machine learning algorithms based on SNS data. *International Journal of Forecasting*, *31*(2), 364-390.

Költringer, C. & Dickinger, A. (2015). Analyzing destination branding and image from online sources: A web content mining approach. *Journal of Business Research*, *68*(9), 1836-1843.

Kontopoulos, E., Berberidis, C., Dergiades, T., & Bassiliades, N. (2013). Ontology-based sentiment analysis of twitter posts. *Expert Systems with Applications*, *40*(10), 4065-4074.

Kou, G. & Lou, C. (2012). Multiple factor hierarchical clustering algorithm for large scale web page and search engine clickstream data. *Annals of Operations Research*, *197*(1), 123-134.

Krishnamoorthy, S. (2015). Linguistic features for review helpfulness prediction. *Expert Systems with Applications*, *42*(7), 3751-3759.

Kurt, D., Inman, J., & Argo, J. (2011). The influence of friends on consumer spending: The role of agency–communion orientation and self-monitoring. *Journal of Marketing Research*, *48*(4), 741-754.

Lau, R. Y., Liao, S. S., Wong, K. F., & Chiu, D. K. (2012). Web 2.0 environmental scanning and adaptive decision support for business mergers and acquisitions. *MIS Quarterly*, *36*(4), 1239-1268.

Laurila, J. K., Gatica-Perez, D., Aad, I., Bornet, O., Do, T. M. T., Dousse, O., Eberle, J., & Miettinen, M. (2012). The mobile data challenge: Big data for mobile computing research. In *Pervasive Computing* (No. EPFL-CONF-192489).

LaValle, S., Lesser, E., Shockley, R., Hopkins, M. S., & Kruschwitz, N. (2011). Big data, analytics and the path from insights to value. *MIT Sloan Management Review*, *52*(2), 20-32.

Lee, C. & Wang, S. (2012). An information fusion approach to integrate image annotation and text mining methods for geographic knowledge discovery. *Expert Systems with Applications*, *39*(10), 8954-8967.

Lee, C., Yang, H., & Wang, S. (2011). An image annotation approach using location references to enhance geographic knowledge discovery. *Expert Systems with Applications*, *38*(11), 13792-13802.

Lee, T. & Bradlow, E. (2011). Automated marketing research using online customer reviews. *Journal of Marketing Research*, *48*(5), 881-894.

Lee, W. (2007). Deploying personalized mobile services in an agent-based environment. *Expert Systems with Applications*, *32*(4), 1194-1207.

Lee, Y., Hosanagar, K., & Tan, Y. (2015). Do I follow my friends or the crowd? Information cascades in online movie ratings. *Management Science*, *61*(9), 2241-2258.

Lenzerini, M. (2002). Data integration: A theoretical perspective. In *Proceedings of the twenty-first ACM SIGMOD-SIGACT-SIGART symposium on Principles of database systems* (pp. 233-246). ACM.

Lew, M. S., Sebe, N., Djeraba, C., & Jain, R. (2006). Content-based multimedia information retrieval: State of the art and challenges. *ACM Transactions on Multimedia Computing, Communications, and Applications, 2*(1), 1-19.

Li, D. & Wang, X. (2017). Dynamic supply chain decisions based on networked sensor data: an application in the chilled food retail chain. *International Journal of Production Research*, *55*(17), 5127-5141.

Li, J., Wang, H., & Bai, X. (2015). An intelligent approach to data extraction and task identification for process mining. *Information Systems Frontiers*, *17*(6), 1195-1208.

Li, K. & Du, T. (2012). Building a targeted mobile advertising system for location-based services. *Decision Support Systems*, *54*(1), 1-8.

Li, N. & Wu, D. (2010). Using text mining and sentiment analysis for online forums hotspot detection and forecast. *Decision Support Systems*, *48*(2), 354-368.

Li, Y., Lin, L., & Chiu, S. (2014). Enhancing targeted advertising with social context endorsement. *International Journal of Electronic Commerce*, *19*(1), 99-128.

Liao, J., Yang, D., Li, T., Wang, J., Qi, Q., & Zhu, X. (2014). A scalable approach for content based image retrieval in cloud datacenter. *Information Systems Frontiers*, *16*(1), 129-141.

Lim, E. P., Chen, H., & Chen, G. (2013). Business intelligence and analytics: Research directions. *ACM Transactions on Management Information Systems, 3*(4), Article 17.

Liu, Z., Yang, P., & Zhang, L. (2013, September). A sketch of big data technologies. In *2013 Seventh International Conference on Internet Computing for Engineering and Science* (pp. 26-29). IEEE.

Lo, S. (2008). Web service quality control based on text mining using support vector machine. *Expert Systems with Applications*, *34*(1), 603-610.

Lu, Y., Jerath, K., & Singh, P. (2013). The emergence of opinion leaders in a networked online community: A dyadic model with time dynamics and a heuristic for fast estimation. *Management Science*, *59*(8), 1783-1799.

Ludwig, S., de Ruyter, K., Friedman, M., Brüggen, E., Wetzels, M., & Pfann, G. (2013). More than words: The influence of affective content and linguistic style matches in online reviews on conversion rates. *Journal of Marketing*, *77*(1), 87-103.

Ludwig, S., De Ruyter, K., Mahr, D., Wetzels, M., Bruggen, E., & De Ruyck, T. (2014). Take their word for it: The symbolic role of linguistic style matches in user communities. *Management Information Systems Quarterly*, *38*(4), 1201-1217.

Lumsdaine, A., Gregor, D., Hendrickson, B., & Berry, J. (2007). Challenges in parallel graph processing. *Parallel Processing Letters*, *17*(01), 5-20.

Luo, C., Wu, F., Sun, J., & Chen, C. W. (2009). Compressive data gathering for large-scale wireless sensor networks. In *Proceedings of the 15th annual international conference on Mobile computing and networking* (pp. 145-156). ACM.

Luo, X., Andrews, M., Fang, Z., & Phang, C. (2014). Mobile targeting. *Management Science*, *60*(7), 1738-1756.

Luo, X., Zhang, J., & Duan, W. (2013). Social media and firm equity value. *Information Systems Research*, *24*(1), 146-163.

Lv, Z., Song, H., Basanta-Val, P., Steed, A., & Jo, M. (2017). Next-generation big data analytics: State of the art, challenges, and future research topics. *IEEE Transactions on Industrial Informatics*, *13*(4), 1891-1899.

Maletic, J. I., & Marcus, A. (2000). Data cleansing: Beyond integrity analysis. In *IQ* (pp. 200-209).

Malewicz, G., Austern, M. H., Bik, A. J., Dehnert, J. C., Horn, I., Leiser, N., & Czajkowski, G. (2010). Pregel: a system for large-scale graph processing. In *Proceedings of the 2010 ACM SIGMOD International Conference on Management of Data* (pp. 135-146). ACM.

Marrese-Taylor, E., Velásquez, J., & Bravo-Marquez, F. (2014). A novel deterministic approach for aspect-based opinion mining in tourism products reviews. *Expert Systems with Applications*, *41*(17), 7764-7775.

Marston, S., Li, Z., Bandyopadhyay, S., Zhang, J., & Ghalsasi, A. (2011). Cloud computing — the business perspective. *Decision Support Systems*, *51*(1), 176-189.

Martínez-Trinidad, J., Beltrán-Martínez, B., & Ruiz-Shulcloper, J. (2000). A tool to discover the main themes in a Spanish or English document. *Expert Systems with Applications*, *19*(4), 319-327.

Mayzlin, D. & Yoganarasimhan, H. (2012). Link to success: How blogs build an audience by promoting rivals. *Management Science*, *58*(9), 1651-1668.

McAfee, A., Brynjolfsson, E., Davenport, T. H., Patil, D. J., & Barton, D. (2012). Big data: The management revolution. *Harvard Business Review*, *90*(10), 61-67.

McCreary, D., & Kelly, A. (2013). *Making sense of NoSQL*. Greenwich: Manning Publications.

Miller, A. & Tucker, C. (2013). Active social media management: The case of health care. *Information Systems Research*, *24*(1), 52-70.

Moe, W. & Trusov, M. (2011). The value of social dynamics in online product ratings forums. *Journal of Marketing Research*, *48*(3), 444-456.

Moniruzzaman, A. B. M., & Hossain, S. A. (2013). Nosql database: New era of databases for big data analytics-classification, characteristics and comparison. *International Journal of Database Theory and Application, 6*(4), 1-14.

Montgomery, A., Li, S., Srinivasan, K., & Liechty, J. (2004). Modeling online browsing and path analysis using clickstream data. *Marketing Science*, *23*(4), 579-595.

Moon, S., Park, Y., & Seog Kim, Y. (2014). The impact of text product reviews on sales. *European Journal of Marketing*, *48*(11/12), 2176-2197.

Moreta, S., & Telea, A. (2007). Multiscale visualization of dynamic software logs. In *Proceedings of the 9th Joint Eurographics/IEEE VGTC conference on Visualization* (pp. 11-18). Eurographics Association.

Moro, S., Cortez, P., & Rita, P. (2015). Business intelligence in banking: A literature analysis from 2002 to 2013 using text mining and latent Dirichlet allocation. *Expert Systems with Applications*, *42*(3), 1314-1324.

Mostafa, M. (2013). More than words: Social networks' text mining for consumer brand sentiments. *Expert Systems with Applications*, *40*(10), 4241-4251.

Mount, M. & Martinez, M. (2014). Social media: A tool for open innovation. *California Management Review*, *56*(4), 124-143.

Musto, C., Semeraro, G., Lops, P., & Gemmis, M. (2015). CrowdPulse: A framework for real-time semantic analysis of social streams. *Information Systems*, *54*, 127-146.

Najafabadi, M. M., Villanustre, F., Khoshgoftaar, T. M., Seliya, N., Wald, R., & Muharemagic, E. (2015). Deep learning applications and challenges in big data analytics. *Journal of Big Data*, *2*(1), 1-21.

Nam, H. & Kannan, P. (2014). The informational value of social tagging networks. *Journal of Marketing*, *78*(4), 21-40.

Nassirtoussi, A., Aghabozorgi, S., Wah, T., & Ngo, D. (2014). Text mining for market prediction: A systematic review. *Expert Systems with Applications*, *41*(16), 7653-7670.

Nassirtoussi, A., Aghabozorgi, S., Wah, T., & Ngo, D. (2015). Text mining of news-headlines for FOREX market prediction: A multi-layer dimension reduction algorithm with semantics and sentiment. *Expert Systems with Applications*, *42*(1), 306-324.

Netzer, O., Feldman, R., Goldenberg, J., & Fresko, M. (2012). Mine your own business: market-structure surveillance through text mining. *Marketing Science*, *31*(3), 521-543.

Neumeyer, L., Robbins, B., Nair, A., & Kesari, A. (2010). S4: Distributed stream computing platform. In *2010 IEEE International Conference on Data Mining Workshops* (pp. 170-177). IEEE.

Ngo-Ye, T. & Sinha, A. (2014). The influence of reviewer engagement characteristics on online review helpfulness: A text regression model. *Decision Support Systems*, *61*, 47-58.

Nguyen, B., Yu, X., Melewar, T., & Chen, J. (2015). Brand innovation and social media: Knowledge acquisition from social media, market orientation, and the moderating role of social media strategic capability. *Industrial Marketing Management*, *51*, 11-25.

Nicholas, D. & Huntington, P. (2003). Micro-mining and segmented log file analysis: A method for enriching the data yield from internet log files. *Journal of Information Science*, *29*(5), 391-404.

Noh, H., Jo, Y., & Lee, S. (2015). Keyword selection and processing strategy for applying text mining to patent analysis. *Expert Systems with Applications*, *42*(9), 4348-4360.

Oestreicher-Singer, G. & Sundararajan, A. (2012). The visible hand? Demand effects of recommendation networks in electronic markets. *Management Science*, *58*(11), 1963-1981.

Okazaki, S. & Taylor, C. (2013). Social media and international advertising: theoretical challenges and future directions. *International Marketing Review*, *30*(1), 56-71.

Ordenes, F., Theodoulidis, B., Burton, J., Gruber, T., & Zaki, M. (2014). Analyzing Customer Experience Feedback Using Text Mining: A Linguistics-Based Approach. *Journal of Service Research*, *17*(3), 278-295.

O'reilly, T. (2007). What is Web 2.0: Design patterns and business models for the next generation of software. *Communications & strategies*, (1), 17.

Orlikowski, W. & Scott, S. (2013). What happens when evaluation goes online? Exploring apparatuses of valuation in the travel sector. *Organization Science*, *25*(3), 868-891.

Özyurt, Ö. & Köse, C. (2010). Chat mining: Automatically determination of chat conversations' topic in Turkish text based chat mediums. *Expert Systems with Applications*, *37*(12), 8705-8710.

Pal, S., Talwar, V., & Mitra, P. (2002). Web mining in soft computing framework: relevance, state of the art and future directions. *IEEE Transactions on Neural Network*, *13*(5), 1163-1177.

Pang, B. & Lee, L. (2008). Opinion Mining and Sentiment Analysis. *Foundations and Trends in Information Retrieval*, *2*(1–2), 1-135.

Paré, G., Trudel, M. C., Jaana, M., & Kitsiou, S. (2015). Synthesizing information systems knowledge: A typology of literature reviews. *Information & Management*, *52*(2), 183-199.

Parhami, B. (2006). *Introduction to parallel processing: algorithms and architectures*. Springer Science & Business Media.

Park, Y. & Chang, K. (2009). Individual and group behavior-based customer profile model for personalized product recommendation. *Expert Systems with Applications*, *36*(2), 1932-1939.

Pedersen, T. & Jensen, C. (2001). Multidimensional database technology. *Computer*, *34*(12), 40-46.

Phillips, F. (2017). A Perspective on 'big data'. *Science and Public Policy*, *4*(5), 730–737

Prates, J., Fritzen, E., Siqueira, S., Braz, M., & de Andrade, L. (2013). Contextual web searches in Facebook using learning materials and discussion messages. *Computers in Human Behavior*, *29*(2), 386-394.

Ramakrishnan, R., & Gehrke, J. (2000). *Database management systems*. McGraw-Hill.

Ransbotham, S., Kane, G., & Lurie, N. (2012). Network characteristics and the value of collaborative user-generated content. *Marketing Science*, *31*(3), 387-405.

Rapp, A., Beitelspacher, L., Grewal, D., & Hughes, D. (2013). Understanding social media effects across seller, retailer, and consumer interactions. *Journal of the Academy of Marketing Science*, *41*(5), 547-566.

Reed, B., Chron, E., Burns, R., & Long, D. (2000). Authenticating network attached storage. *IEEE Micro*, *20*(1), 49-57.

Roberts, C. (2006). Radio frequency identification (RFID). *Computers & Security*, *25*(1), 18-26.

Roosta, S. H. (2012). *Parallel processing and parallel algorithms: theory and computation*. Springer Science & Business Media.

Roth, P., Bobko, P., Van Iddekinge, C., & Thatcher, J. (2013). Social media in employee-selection-related decisions: A research agenda for uncharted territory. *Journal of Management*, *42*(1), 269-298.

Russom, P. (2011). Big data analytics. *TDWI Best Practices Report, Fourth Quarter*, 1-35.

Russom, P. (2012). Analytic databases for big data. *TDWI Checklist Report*, TDWI Research.

Sabnis, G. & Grewal, R. (2015). Cable news wars on the internet: Competition and user-generated content. *Information Systems Research*, *26*(2), 301-319.

Sagiroglu, S., & Sinanc, D. (2013). Big data: A review. In *Collaboration Technologies and Systems (CTS), 2013 International Conference on* (pp. 42-47). IEEE.

Salihoglu, S., & Widom, J. (2013). GPS: a graph processing system. In *Proceedings of the 25th International Conference on Scientific and Statistical Database Management* (pp. 22). ACM.

Sarawagi, S., & Bhamidipaty, A. (2002). Interactive deduplication using active learning. In *Proceedings of the eighth ACM SIGKDD international conference on Knowledge discovery and data mining* (pp. 269-278). ACM.

Schäfer, K. & Kummer, T. (2013). Determining the performance of website-based relationship marketing. *Expert Systems with Applications*, *40*(18), 7571-7578.

Schniederjans, D., Cao, E., & Schniederjans, M. (2013). Enhancing financial performance with social media: An impression management perspective. *Decision Support Systems*, *55*(4), 911-918.

Schweidel, D. & Moe, W. (2014). Listening in on social media: A joint model of sentiment and venue format choice. *Journal of Marketing Research*, *51*(4), 387-402.

Seuring, S., & Gold, S. (2012). Conducting content-analysis based literature reviews in supply chain management. *Supply Chain Management: An International Journal, 17*(5), 544-555.

Shahabi, C. & Banaei-Kashani, F. (2003). Efficient and anonymous web-usage mining for web personalization. *INFORMS Journal on Computing*, *15*(2), 123-147.

Sheng, J., Amankwah-Amoah, J., & Wang, X. (2017). A multidisciplinary perspective of big data in management research. *International Journal of Production Economics*, *191*, 97–112

Shieh, W. (2011). OFDM for flexible high-speed optical networks. *Journal of Lightwave Technology*, *29*(10), 1560-1577.

Shiroishi, Y., Fukuda, K., Tagawa, I., Iwasaki, H., Takenoiri, S., & Tanaka, H. et al. (2009). Future options for HDD storage. *IEEE Transactions on Magnetics*, *45*(10), 3816-3822.

Short, J. (2009). The art of writing a review article. *Journal of Management*, *35*(6), 1312-1317.

Short, J., Ketchen, D., Shook, C., & Ireland, R. (2009). The concept of "opportunity" in entrepreneurship research: Past accomplishments and future challenges. *Journal of Management*, *36*(1), 40-65.

Shriver, S., Nair, H., & Hofstetter, R. (2013). Social ties and user-generated content: Evidence from an online social network. *Management Science*, *59*(6), 1425-1443.

Shvachko, K., Kuang, H., Radia, S., & Chansler, R. (2010). The hadoop distributed file system. In *2010 IEEE 26th symposium on mass storage systems and technologies* (pp. 1-10). IEEE.

Singh, J. (2014). Big data analytic and mining with machine learning algorithm. *International Journal of Information and Computation Technology*, *4*(1), 33-40.

Singh, P., Sahoo, N., & Mukhopadhyay, T. (2014). How to attract and retain readers in enterprise blogging?. *Information Systems Research*, *25*(1), 35-52.

Singh, S., Hillmer, S., & Wang, Z. (2011). Efficient methods for sampling responses from large-scale qualitative data. *Marketing Science*, *30*(3), 532-549.

Sonnier, G., McAlister, L., & Rutz, O. (2011). A dynamic model of the effect of online communications on firm sales. *Marketing Science*, *30*(4), 702-716.

Sridhar, S. & Srinivasan, R. (2012). Social influence effects in online product ratings. *Journal of Marketing*, *76*(5), 70-88.

Stonebraker, M., Çetintemel, U., & Zdonik, S. (2005). The 8 requirements of real-time stream processing. *ACM SIGMOD Record*, *34*(4), 42-47.

Suh, J., Park, C., & Jeon, S. (2010). Applying text and data mining techniques to forecasting the trend of petitions filed to e-People. *Expert Systems with Applications*, *37*(10), 7255-7268.

Sun, M. (2012). How does the variance of product ratings matter?. *Management Science*, *58*(4), 696-707.

Suneetha, K. R., & Krishnamoorthi, R. (2009). Identifying user behavior by analyzing web server access log file. *IJCSNS International Journal of Computer Science and Network Security*, *9*(4), 327-332.

Sunikka, A. & Bragge, J. (2012). Applying text-mining to personalization and customization research literature – Who, what and where?. *Expert Systems with Applications*, *39*(11), 10049-10058.

Talia, D. (2013). Clouds for scalable big data analytics. *Computer, 46*(5), 98-101.

Tang, C. & Guo, L. (2015). Digging for gold with a simple tool: Validating text mining in studying electronic word-of-mouth (eWOM) communication. *Marketing Letters*, *26*(1), 67-80.

Telikepalli, R., Drwiega, T., & Yan, J. (2004). Storage area network extension solutions and their performance assessment. *IEEE Communications Magazine*, *42*(4), 56-63.

Thelwall, M. (2001a). A web crawler design for data mining. *Journal of Information Science*, *27*(5), 319-325.

Thelwall, M. (2001b). Web log file analysis: backlinks and queries. *Aslib Proceedings*, *53*(6), 217-223.

Thorleuchter, D. & Van den Poel, D. (2012). Predicting e-commerce company success by mining the text of its publicly-accessible website. *Expert Systems with Applications*, *39*(17), 13026-13034.

Thorleuchter, D. & Van den Poel, D. (2013). Web mining based extraction of problem solution ideas. *Expert Systems with Applications*, *40*(10), 3961-3969.

Thorleuchter, D., van den Poel, D., & Prinzie, A. (2010). Mining ideas from textual information. *Expert Systems with Applications*, *37*(10), 7182-7188.

Thorleuchter, D., van den Poel, D., & Prinzie, A. (2012). Analyzing existing customers' websites to improve the customer acquisition process as well as the profitability prediction in B-to-B marketing. *Expert Systems with Applications*, *39*(3), 2597-2605.

Tirunillai, S. & Tellis, G. (2014). Mining marketing meaning from online chatter: Strategic brand analysis of big data using Latent Dirichlet Allocation. *Journal of Marketing Research*, *51*(4), 463-479.

Tsai, C. W., Lai, C. F., Chao, H. C., & Vasilakos, A. V. (2016). Big data analytics. In *Big Data Technologies and Applications* (pp. 13-52). Springer.

Tsai, T. & Lin, C. (2012). Exploring contextual redundancy in improving object-based video coding for video sensor networks surveillance. *IEEE Transactions on Multimedia*, *14*(3), 669-682.

Turban, E., Sharda, R., Aronson, J. E., & King, D. (2008). *Business intelligence: A managerial approach*. Upper Saddle River, NJ: Pearson Prentice Hall.

Ucbasaran, D., Shepherd, D., Lockett, A., & Lyon, S. (2013). Life after business failure: The process and consequences of business failure for entrepreneurs. *Journal of Management*, *39*(1), 163-202.

Ur-Rahman, N. & Harding, J. (2012). Textual data mining for industrial knowledge management and text classification: A business oriented approach. *Expert Systems with Applications*, *39*(5), 4729-4739.

Van Iddekinge, C., Lanivich, S., Roth, P., & Junco, E. (2013). Social media for selection? Validity and adverse impact potential of a Facebook-based assessment. *Journal of Management, 42*(7), 1811–1835.

Wamba, S. F., Gunasekaran, A., Akter, S., Ren, S. J. F., Dubey, R., & Childe, S. J. (2017). Big data analytics and firm performance: Effects of dynamic capabilities. *Journal of Business Research*, *70*, 356-365.

Wang, C., Lu, J., & Zhang, G. (2007). Mining key information of web pages: A method and its application. *Expert Systems with Applications*, *33*(2), 425-433.

Wang, D., Zhu, S., & Li, T. (2013). SumView: A Web-based engine for summarizing product reviews and customer opinions. *Expert Systems with Applications*, *40*(1), 27-33.

Wang, F. & Liu, J. (2011). Networked wireless sensor data collection: Issues, challenges, and approaches. *IEEE Communications Surveys & Tutorials*, *13*(4), 673-687.

Wang, K., Ting, I., & Wu, H. (2013). Discovering interest groups for marketing in virtual communities: An integrated approach. *Journal of Business Research*, *66*(9), 1360-1366.

Wang, M., Ni, B., Hua, X. S., & Chua, T. S. (2012). Assistive tagging: A survey of multimedia tagging with human-computer joint exploration. *ACM Computing Surveys (CSUR)*, *44*(4), Article No. 25.

Wang, X., Mai, F., & Chiang, R. (2013). Database submission —Market dynamics and user-generated content about tablet computers. *Marketing Science*, *33*(3), 449-458.

Wang, Y., Kung, L., & Byrd, T. A. (2018). Big data analytics: Understanding its capabilities and potential benefits for healthcare organizations. *Technological Forecasting and Social Change, 126*, 3-13.

Watson, H. & Wixom, B. (2007). The current state of business intelligence. *Computer*, *40*(9), 96-99.

Watson, H. J. (2014). Tutorial: Big data analytics: Concepts, technologies, and applications. *Communications of the Association for Information Systems*, *34*(1), 1247-1268.

Webster, J., & Watson, R. T. (2002). Analyzing the past to prepare for the future: Writing a literature review. *MIS Quarterly*, xiii-xxiii.

Wei, C., Chiang, R., & Wu, C. (2006). Accommodating individual preferences in the categorization of documents: A personalized clustering approach. *Journal of Management Information Systems*, *23*(2), 173-201.

Wei, C., Hu, P., Tai, C., Huang, C., & Yang, C. (2007). Managing word mismatch problems in information retrieval: A topic-based query expansion approach. *Journal of Management Information Systems*, *24*(3), 269-295.

Wei, C., Yang, C., & Hsiao, H. (2008). A collaborative filtering-based approach to personalized document clustering. *Decision Support Systems*, *45*(3), 413-428.

Wei, C., Yang, C., & Lin, C. (2008). A Latent Semantic Indexing-based approach to multilingual document clustering. *Decision Support Systems*, *45*(3), 606-620.

Weng, S. & Liu, C. (2004). Using text classification and multiple concepts to answer e-mails. *Expert Systems with Applications*, *26*(4), 529-543.

Witten, I. H., Frank, E., & Hall, M. A. (2011). *Data mining: Practical machine learning tools and techniques* (3rd ed.). Morgan Kaufmann.

Wu, L. (2013). Social network effects on productivity and job security: Evidence from the adoption of a social networking tool. *Information Systems Research*, *24*(1), 30-51.

Wu, X., Kumar, V., Ross Quinlan, J., Ghosh, J., Yang, Q., & Motoda, H. et al. (2008). Top 10 algorithms in data mining. *Knowledge and Information Systems*, *14*(1), 1-37.

Xu, J., Forman, C., Kim, J., & Van Ittersum, K. (2014). News media channels: Complements or substitutes? Evidence from mobile phone usage. *Journal of Marketing*, *78*(4), 97-112.

Yang, H. & Lee, C. (2004). A text mining approach on automatic generation of web directories and hierarchies. *Expert Systems with Applications*, *27*(4), 645-663.

Yang, H. & Lee, C. (2008). Image semantics discovery from web pages for semantic-based image retrieval using self-organizing maps. *Expert Systems with Applications*, *34*(1), 266-279.

Yang, H. (2009). Automatic generation of semantically enriched web pages by a text mining approach. *Expert Systems with Applications*, *36*(6), 9709-9718.

Yang, W., Cheng, H., & Dia, J. (2008). A location-aware recommender system for mobile shopping environments. *Expert Systems with Applications*, *34*(1), 437-445.

Ye, Q., Zhang, Z., & Law, R. (2009). Sentiment classification of online reviews to travel destinations by supervised machine learning approaches. *Expert Systems with Applications*, *36*(3), 6527-6535.

Yeh, I., Lien, C., Ting, T., & Liu, C. (2009). Applications of web mining for marketing of online bookstores. *Expert Systems with Applications*, *36*(8), 11249-11256.

Yoon, J. (2012). Detecting weak signals for long-term business opportunities using text mining of Web news. *Expert Systems with Applications*, *39*(16), 12543-12550.

You, K., Dal Bianco, S., Lin, Z., & Amankwah-Amoah, J. (2018). Bridging technology divide to improve business environment: Insights from African nations. Journal of Business Research.

Yu, Y., Duan, W., & Cao, Q. (2013). The impact of social and conventional media on firm equity value: A sentiment analysis approach. *Decision Support Systems*, *55*(4), 919-926.

Zeng, J., Wu, C., & Wang, W. (2010). Multi-grain hierarchical topic extraction algorithm for text mining. *Expert Systems with Applications*, *37*(4), 3202-3208.

Zhan, J., Loh, H., & Liu, Y. (2009). Gather customer concerns from online product reviews – A text summarization approach. *Expert Systems with Applications*, *36*(2), 2107-2115.

Zhang, X., Li, S., Burke, R., & Leykin, A. (2014). An examination of social influence on shopper behavior using video tracking data. *Journal of Marketing*, *78*(5), 24-41.

Zhang, Y. & Jiao, J. (2007). An associative classification-based recommendation system for personalization in B2C e-commerce applications. *Expert Systems with Applications*, *33*(2), 357-367.

Zissis, D. & Lekkas, D. (2011). Securing e-Government and e-Voting with an open cloud computing architecture. *Government Information Quarterly*, *28*(2), 239-251.
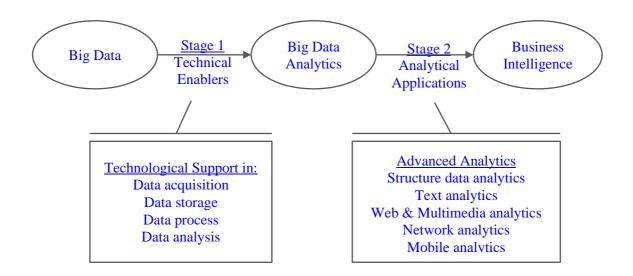
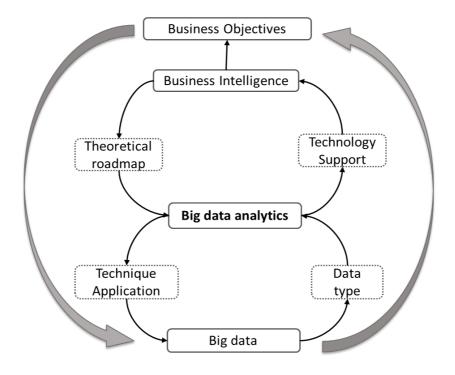**Figure 1. Road Map to Review Studies on Big Data, BDA and Business Intelligence**

Figure 2. Framework for big data research

Table 1. List of Studies on Web and Mobile Analytics

	Analytics	Studies	Key area
Web	*Web mining*	Baek et al. (2012)	Online reviews
		Castellanos et al. (2012)	Streaming data analytics
		Chau and Chen (2008)	Web page classification
		Chau and Xu (2007)	Online community
		Chung et al. (2005)	Knowledge discovery
		Costa et al. (2012)	E-commerce
		D'Haen et al. (2013)	Predict profitability
		Ding et al. (2015)	Real-time analysis
		Ho et al. (2011)	Personalization
		Költringer and Dickinger (2015)	Destination brand
		Shahabi and Banaei-Kashani (2003)	Personalization
		Thorleuchter and Van den Poel (2013)	Idea mining
		Wang et al. (2007)	Web content mining
		Yeh et al. (2009)	Customer prediction
		Sonnier et al. (2011)	Online communication
	Web analytics	Järvinen and Karjaluoto (2015)	Digital marketing
		Kou and Lou (2012)	Web search engines
		Lau et al. (2012)	International investment
	Clickstream	Chatterjee et al. (2003)	Online advertising
		Montgomery et al. (2004)	Personalization
		Schäfer and Kummer (2013)	E-commerce
		Huang and Mieghem (2014)	Inventory management
	Cloud computing	Marston et al. (2011)	Business-related issues
		Delen and Demirkan (2013)	Service orientation
		Guo et al. (2014)	Apparel manufacturing
		Zissis and Lekkas (2011)	E-government
Multimedia	*Image retrieval and annotation*	Lee and Wang (2012)	Geographic knowledge discovery
		Lee et al. (2011)	Analytics technique
		Liao et al. (2014)	Analytics technique
		Yang and Lee (2008)	Image semantics
	Video data	Zhang et al. (2014)	Social influence
Mobile	*Mobile sensing*	Andrews et al. (2016)	Mobile advertising
		Fong et al. (2015)	Mobile targeting
		Ghose et al. (2012)	Mobile Internet usage
		Lee (2007)	Personalization
		Li and Du (2012)	Mobile advertising
		Li and Wang (2015)	Food supply chain
		Luo et al. (2014)	Mobile targeting
		Yang et al. (2008)	Recommendation
	Mobile network	Chung et al. (2015)	Personalization
		Ghose and Han (2011)	Mobile Internet usage
		Wang et al. (2013)	Market dynamics
	Mobile app	Ghose and Han (2014)	Mobile apps demand
		Xu et al. (2014)	Online media

Table 2. Unanswered Questions for Future Research

Area	Key theme	Unanswered questions
Big data	Structured data	1. To what extent can conventional techniques be applied to big structured data? 2. How can structured data be effectively combined with unstructured data to design business strategy?
	Unstructured data	3. How can unstructured data be effectively utilised to address business problems? 4. What is the most effective method to concurrently analyse both structured and unstructured data?
Big data analytics	Technique improvement	5. To what extent does an advanced technique improve the effectiveness of analytics? 6. What is the most effective mechanism to assess and compare the effectiveness of different analytics techniques?
	Technique application	7. How best to transfer and apply techniques in computer science, engineering and other technological subjects into analysing big data for general management?
Business intelligence	Theoretical road map	8. What are the main steps for adopting big data in enterprise business intelligence process? 9. What are influential paths and mechanisms for value creation in organizations?
	Technological support	10. How best to incorporate advanced analytics techniques into current business intelligence system to improve BDA ability? 11. What is the most appropriate business intelligence platform for firms?

Appendix A. List of studies on the topic of text analytics

Studies	Key area/application	Key findings
Aliguliyev (2009a)	Document clustering	By signing weights, document clustering can be improved and thus the performance of search engines.
Aliguliyev (2009b)	Document summarization	The sentence-clustering based method improves document summarization performance.
Archak et al. (2011)	Consumers reviews	The textual content of product reviews has a significant impact on consumers' choices, which can be used for predicting future sales.
Balakrishnan et al. (2010)	Capital market	Narrative disclosure has value-relevant information to predict market performance.
Bao and Datta (2014)	Risk type evaluation	Risk types can be measured with corporate textual data of risk disclosure.
Baralis et al. (2013)	Document summarization	The proposed method can effectively perform ontology-based analysis of document sentences.
Beebe et al. (2011)	Analytics technique	The extended approach to digital forensic text string search process improves information retrieval effectiveness.
Cao et al. (2011)	Online reviews	More helpfulness is found in reviews with extreme opinions than in those with mixed or neutral opinions.
Chen and Tseng (2011)	Quality evaluation	The proposed method can effectively evaluate review quality and thus improve opinion mining efficiency.
Chou et al. (2010)	Analytics technique	Combining filter and wrapper attribute selection approaches provides more efficient and accurate classification results.
Chung and Tseng (2012)	Product review	The proposed framework can effectively detect the relationship between reviews and ratings.
Colace et al. (2014)	Analytics technique	A single label text classification method performs better when labelled examples are limited.
Coussement and van den Poel (2008)	Churn prediction	Textual information in emails is beneficial to improve churn prediction performance.
Crawford Camiciottoli et al. (2014)	Brand associations	Consistent brand associations are detected in an online community of international consumers.
Delen and Crossland (2008)	Literature survey	Text mining is of great use to analyse textual information and identify patterns of research topics.
Duan et al. (2011)	Analytics technique	This is an effective approach to extract bilingual multiword expression.
Fan et al. (2006)	Personalization	Online information routing of personalised information can be improved with the two-stage model.
Fuller et al. (2011)	Deception detection	Text and data mining are combined to automatically detect deception with better accuracy.
Glancy and Yadav (2011)	Fraud detection	A computational fraud detection model on textual data is effective for detecting financial fraud in reports.
Guerreiro et al. (2016)	Cause-related marketing	Text mining techniques can help review literature and uncover topics.
Hashimi et al. (2015)	Selection criteria	A set of criteria is proposed to evaluate the effectiveness of different text mining techniques.
He (2013a)	Online interaction	Video stream data are used to identify patterns in online learning behaviours.

He (2013b)	Case-based reasoning	Text mining and Web 2.0 tools can be beneficial to CBR systems with a better user experience.
Hu et al. (2012)	Online reviews	Manipulation by firms is found in online reviews, particularly in product ratings.
Hyung et al. (2014)	Music recommendation	Analysing listeners' text from a radio station's online bulletin board is beneficial for recommending music.
Janasik et al. (2009)	Research method	SOM enables better effectiveness of text mining to improve inference quality in qualitative research.
Kayser and Blind (2016)	Foresight practice	Text mining contributes to foresight by broadening the knowledge base.
Krishnamoorthy (2015)	Online reviews	The proposed model based on linguistic features can achieve better predictions of online review helpfulness.
Lee and Bradlow (2011)	Market structure analysis	The proposed system automatically processes text from online reviews and improves marketing strategy.
Li et al. (2015)	Process mining	The proposed method improves data extraction and task identification from event logs.
Lo (2008)	Web service quality	Text analysis helps in classifying customers' messages and identifying service quality.
Ludwig et al. (2013)	Online reviews	Powerful reviews should be identified, encouraged and promoted with typical linguistic style.
Ludwig et al. (2014)	Community identification	Linguistic style match in user communities indicates community identification and fosters greater participation quantity and quality.
Martínez-Trinidad et al. (2000)	Analytics technique	The tool is used for detecting themes in documents. High co-occurrence between two concepts implies strong relationships.
Moon et al. (2014)	Product review	Analysing text of consumers' product reviews can enhance sales.
Moro et al. (2015)	BI in banking	Text mining is used to review the literature on business intelligence application in banking.
Mostafa et al. (2013)	Brand sentiment	Tweets are used to detect consumers' sentiments, and a general positive attitude is found in the sample.
Nassirtoussi et al. (2014)	Market prediction	The review identifies a clear frame of discussion on market prediction using online text mining.
Nassirtoussi et al. (2015)	Market prediction	A multi-layer algorithm integrates sentiment analysis to tackle textual information with a focus on financial market prediction.
Netzer et al. (2012)	Market structure analysis	Online user-generated contents can improve mapping of market structure.
Ngo-Ye and Sinha (2014)	Online reviews	Review text and reviewer engagement characteristics can predict the helpfulness of online reviews.
Noh et al. (2015)	Patent analysis	Keyword strategies of text mining can be applied for patent analysis with better validity and reliability.
Ordenes et al. (2014)	Customer feedback	The advanced linguistic-based text mining model can more deeply analyse customers' experiences.
Özyurt and Köse (2010)	Chat mining	Mining chat conversations using proposed classification can effectively detect topics.
Singh et al. (2014)	Employee blogs	Textual characteristics of blogs affect readers' attention and retention, which can help improve understanding of reading behaviour in communities.

Suh et al. (2010)	Petition trend forecast	Text and data mining are combined to efficiently detect and forecast trend of the petition in e-government.
Sunikka and Bragge (2012)	Research profiling	Text mining is applied to review the literature on personalization to identify future research opportunity.
Tang and Guo (2015)	Electronic word-of-mouth	Linguistic indicators from text can predict word-of-mouth attitudes about products and services.
Thorleuchter and van den Poel (2012)	E-commerce	Textual information on companies' websites is useful for predicting commercial success.
Thorleuchter et al. (2010)	Idea mining	An idea mining approach is proposed to automatically discover new ideas from textual information.
Thorleuchter et al. (2012)	Profitability prediction	Current customers' website information can be used to identify future profitable customers more accurately.
Ur-Rahman and Harding (2012)	Analytics technique	Classification accuracies are improved by classifying textual data into different classes.
Wang et al. (2013)	Online reviews	A web-based system can automatically extract and summarise information from review documents,
Wei et al. (2006)	Document clustering	The personalised document-clustering approach can achieve better clustering effectiveness.
Wei et al. (2007)	Query expansion	A topic-based method for query expansion is more effective for addressing word mismatch problem in information retrieval.
Wei et al. (2008)	Knowledge map	It is effective to generate knowledge maps using the multilingual document clustering method.
Wei et al. (2008)	Personalization	Better effectiveness and personalisation are achieved by the extended document-clustering technique.
Weng and Liu (2004)	E-mail responding	Different concepts are integrated into classification to effectively extract information and reply e-mails.
Yang (2009)	Web page annotation	The proposed methods can automatically generate metadata for the web page for semantic analysis.
Yang and Lee (2004)	Web directories	The corpus-based method automatically and efficiently illustrates web directory hierarchy with labels.
Yoon (2012)	Weak signal detection	Keyword-based text mining is useful for identifying weak signal topics for future business planning.
Zeng et al. (2010)	Analytics technique	A multi-grain hierarchical topic structure that can provide a description for subtopics outperforms other methods.
Zhan et al. (2009)	Online reviews	Text in online product reviews can be automatically summarised based on internal topic structure.
Zhang and Jiao (2007)	Recommendation and e-commerce	The proposed associative classification-based recommendation system can be applied for personalization in B2C e-commerce application.

Appendix B. List of studies on network analytics

Studies	Key area/application	Key findings
Social media analysis		
Abrahams et al. (2014)	Quality management	Social media content and text are of great use for product defect discovery.
Beverungen et al. (2015)	Free labour	Social media users working as free labour challenges Marxist organization studies.
Chan et al. (2017)	Operation management	Social media comments are analysed to discover useful information for operations management.
Chan et al. (2016)	Analytics technique	A mixed approach to analysing social media data can enhance operational decision-making.
Claussen et al. (2013)	Mobile apps success	App success relies on rewarding users' engagement, high user ratings and frequent updating.
Feuls et al. (2014)	Unemployment	Use of social media can help unemployed people to maintain and cultivate social networks.
Gao et al. (2015)	Emotion analysis	A rule-based approach is proposed to detect emotion-causing component from social media data.
Ghose et al. (2012)	Ranking system	The proposed ranking system analyses user-generated content to assess customers' preferences and can provide best-value hotels.
Gopinath et al. (2013)	Blogs and advertising	The effects of pre- and post-release blog and advertising vary across markets with different demographic characteristics and groups.
He et al. (2015)	Marketing	Competitive social media analysis with sentiment benchmarks is beneficial to marketing.
Iyer and Katona (2016)	Marketing	The incentives for social communication in social media decrease with an increased span of communications.
Jang et al. (2013)	Deep sentiment analysis	The proposed method can determine customers' value structure and causality for identifying a niche market.
Kim et al. (2015)	Forecasting earning	Social network information and a machine learning algorithm can improve forecasting accuracy.
Li et al. (2014)	Targeted advertising	The effectiveness of advertisement can be enhanced by leveraging social context and social influence.
Luo et al. (2013)	Firm equity value	Social media metrics significantly indicate firms' equity value with stronger and faster predictive relationships than traditional online behavioural metrics.
Mayzlin and Yoganarasimhan (2012)	Web logs linking	Blogs with links to other blogs signal as high news-breaking ability and lead to more readers and enhanced learning.
Miller and Tucker (2013)	Business value of IT	Social media management in health-care sector leads to more user-generated content by employees, and it may have adverse effects.

Moe and Trusov (2011)	Online product rating	Online product ratings dynamics have direct and immediate effects on sales and indirect impact on further ratings.
Mount and Martinez (2014)	Open innovation	Social media has benefits to be applied to the open innovation process.
Nam and Kannan (2014)	Brand performance	Social tagging has great implications for brand performance measurement and brand equity management.
Nguyen et al. (2015)	Brand innovation	Social media strategic capability can enhance brand innovation and moderate between innovation, knowledge acquisition and market orientation.
Oestreicher-Singer and Sundararajan (2012)	Recommendation	Demands increase with explicit visibility of a co-purchase relationship in a recommendation network.
Okazaki and Taylor (2013)	International advertising	Three theoretical foundations are identified for future research on social media use for international advertising.
Orlikowski and Scott (2014)	Online evaluation	Online and traditional valuation is significantly different in performativity, which has organizational implications.
Prates et al. (2013)	Web search engines	Social media data are adopted to improve contextual information extraction through web searching.
Rapp et al. (2013)	Contagion effect	Positive contagion effects of social media use are found in enhancing brand performance, retailer performance and consumer-retailer loyalty.
Roth et al. (2013)	Personnel decision	The use of social media in human resources practice has great importance for organizations, individuals and society and needs further study.
Sabnis and Grewal (2015)	Competition	There is a significant relationship between competitor user-generated content and firms' performance.
Schniederjans et al. (2013)	Impression management	Social media has a positive impact on impression management and can influence firms' financial performance.
Schweidel and Moe (2014)	Brand sentiment	The analysis of brand sentiment cannot ignore the differences across different social media venue formats.
Shriver et al. (2013)	Social network analysis	Online user-generated content has a positive relationship with their social ties, and it has network effects that boost advertising and revenue growth.
Singh et al. (2011)	Analytics technique	The proposed method of sampling qualitative comments improves the effectiveness of text mining.
Sun (2012)	Product rating	A higher variance of product ratings helps with sales increase if and only if the average rating is low.
Tirunillai and Tellis (2014)	Brand performance	Dynamic analysis of online user-generated content can reflect consumers' satisfaction with the quality and thus improve competitive brand positions.
Van Iddekinge et al. (2013)	Personnel decision	Social media information of job applicants is irrelevant and invalid to recruitment selection.

Wu (2013)	Social network effect	Social media can enrich network information, which has a positive effect on work productivity and job security.
Social network analysis		
Autio et al. (2013)	Entrepreneurial action	Online user community information about users' needs can stimulate entrepreneurial action.
Fang et al. (2013)	Adoption probabilities	The method can effectively predict adoption probabilities based on key factors that affect adoption decisions.
Garg et al. (2011)	Information diffusion	Information diffusion and discovery in online social networks are effectively measured using proposed methods.
Ransbotham et al. (2012)	Information value	The value of collaborating user-generated content depends on contributors' efforts as well as their network.
Wang et al. (2013)	E-commerce	Social network analysis and web mining are integrated to detect groups in virtual communities for better recommendations and marketing.
Community detection		
Feng et al. (2015)	Personalization	A time-weighted overlapping community detection method performs better to predict users' interests and thus give personalised recommendations.
Johnson et al. (2015)	Leadership	Use of language shapes online community dynamics. Similar use is found from leaders and other participants.
Social influence		
Berger (2014)	Word of mouth	Word of mouth is goal-driven and has five key functions with psychological factors behind them.
Cascio et al. (2015)	Consumer behaviour	People tend to change their initial recommendation decisions to be consistent with peer recommendations.
Cheng and Ho (2015)	Online reviews	There is a positive impact of reviewers' number of followers, level of expertise, image count and word count on readers' perceptions of reviews.
Eisingerich et al. (2015)	Word of mouth	Consumers' willingness to engage in word-of-mouth on online social sites is lower than face-to-face WOM.
Goes et al. (2014)	Online community	Popular users in online community tend to generate more objective, negative and varied product reviews.
Goh et al. (2013)	Brand community	Users' and marketers' engagement in social media brand communities has a positive impact on purchase behaviour and expenditures.
Haenlein (2011)	Customer relations	There is a strong positive social network relationship in customer-level revenue.
Hennig-Thurau et al. (2015)	Microblogging	A microblog containing post-purchase quality evaluation information affects early movie adoption behaviours.
Hildebrand et al. (2013)	Consumer satisfaction	Feedback on product features from other community members has a negative influence on customers' satisfaction with self-designed products.

Kurt et al. (2011)	Consumer behaviour	Agency-oriented consumers spend more when shopping with friends, while communion-oriented consumers do not.
Lee et al. (2015)	Online rating	Prior ratings by friends positively influence users' ratings, while social networking decreases possible effects on ratings by the crowd.
Lu et al. (2013)	Opinion leadership	The activity of the online community and style of writing the review are strong drivers for network growth.
Sridhar and Srinivasan (2012)	Online ratings	The online rating has a social influence on other consumers that is contingent on product experience.
Sentiment analysis		
Alfaro et al. (2013)	Opinion mining	Opinion trends can be detected from weblog comments, providing useful information for decision-making.
Balahur et al. (2012)	Emotion detection	When words have no affective meaning, an approach using EmotiNet is more valid for emotion detection.
Colace et al. (2015)	Recommendation	User-generated data in the online social network can support customised recommendations.
Colace et al. (2015)	Analytics technique	The proposed method uses a mixed graph of terms that yields more effective sentiment classification.
Da Silva et al. (2014)	Microblogging	Classifier ensembles can improve classification accuracy of microblogging sentiment analysis.
Das and Chen (2007)	Sentiment extraction	Small investor sentiment from stock message boards is related to stock value and has an impact on investors' opinions and financial management.
Dehkharghani et al. (2014)	Causal rule discovery	Sentiment causal rules effectively summarise important relationships and sentiment from social media textual data.
Deng et al. (2014)	Term weighting	The supervised term-weighting approach gives more accurate results in sentiment analysis.
Fang et al. (2014)	Sentiment classification	The proposed approach with sentiment information from source-domain labelled data and preselected sentiment works is proved efficient.
Fersini et al. (2014)	Polarity classification	Ensemble learning is introduced to predict polarity with better accuracy.
García-Cumbreras et al. (2013)	Collaborative filtering	Sentiment analysis incorporated in collaborative filtering algorithms improves rating prediction and recommendation.
García-Moya et al. (2013)	Analytics technique	The proposed system allows analysing sentiment data in the corporate data warehouse.
Gopaldas (2014)	Consumer sentiment	The proposed theory of marketplace sentiments advances studies on consumers with a sociocultural perspective.
Homburg et al. (2015)	Marketing performance	Firms' active participation in the online community among consumers has a negative impact on consumers' sentiment and returns.
Kang and Park (2014)	Customer satisfaction	Customers' satisfaction can be measured by customers' reviews with greater effectiveness and efficiency.

Khan et al. (2014)	Microblogging	The proposed hybrid approach can achieve higher accuracy in sentiment classification.
Kontopoulos et al. (2013)	Microblogging	An ontology-based method is more efficient for analysing opinions towards specific topics in microblogs.
Li and Wu (2010)	Forums hotspot detection	The proposed method can improve hotspot detection and forecast from online forums.
Marrese-Taylor et al. (2014)	Opinion mining	The aspect-based opinion mining approach can be used in tourism product reviews with extension.
Musto et al. (2015)	Social streams	The proposed framework can effectively process semantic analysis of textual content in social streams.
Ye et al. (2009)	Sentiment classification	Well-trained machine learning algorithms can classify sentiment polarities of reviews with high accuracy.
Yu et al. (2013)	Firm equity value	Social media has greater effects on firm stock performance than conventional media.