Turkish Online Journal of Qualitative Inquiry (TOJQI)
Volume 12, Issue 6, July, 2021: 8074-8099

Big Data in the Transport Industry - An Indispensable Trend Approaching Logistics 4.0

Hoang Phuong Nguyen[1], Viet Duc Bui[2]

Abstract

The transportation and logistics industries are making very important changes, with many advanced technology solutions being launched. The increasing development of technology and the application of information technology in the logistics industry have made it possible for us to collect a huge amount of data from many different sources. Enterprises tend to use Big data for analysis during peak waves to make it easier to exchange and share information from that data. Using Big data in supply chain management will help businesses manage and forecast customers' needs more accurately. Moreover, it better understands the shopper cycle and preferences for better customer care. In the logistics industry, Big data and data analysis help shipping companies make the most accurate decisions, but the challenge of Big data is security and safety control. In these large data sets of businesses, there will be personal information about their customers. Therefore, this article presents a comprehensive assessment of the opportunities and challenges that Big data has been bringing to the logistics industry, especially the trend of logistics 4.0. Based on a general assessment of the concept and characteristics of Big data and logistics 4.0, this work also focuses on clarifying the knowledge domain about potential applications of Big data in logistics 4.0.

Keywords: *Logistics, transportation, big data, industry 4.0.*

Introduction

Nowadays, innovation is associated with modernity and advancement in every field including society, educational system, economy, and economic science. Innovation means finding resolutions for competitive advantage, which contributes to the increase in the development of the economy and society as well as enhances the quality of life. The fourth industrial revolution (so-called Industry 4.0), which is a result of the fast development of emerging technology, mentions the optimization and digitalization of industrial processes with new

[1] Academy of Politics Region II, Ho Chi Minh City, Vietnam, nghoangphuong11@gmail.com

[2] Ho Chi Minh city University of Technology (HUTECH), Ho Chi Minh city, Vietnam, bv.duc@hutech.edu.vn

technologies being used including Cloud Computing, Artificial Intelligence, IoT, and Big Data [1][2]. Despite the presence of the Industry 4.0 concept for a while, just approximately 48 percent of producing factories are ready to cope with changes in technology which is supported by these building blocks [3]. Data are being formed on a large scale ever before thanks to the progressions of information technologies as well as the immense increase in international data resources, which results in the source of data with a big amount and increasing growth. Thus, the vast generation of data, and the new possibilities for exploration along with analysis and management challenges, have created a new term, known as Big Data.

Following shifts in retail sales, logistics has changed [4]. According to Wang et al. [5], there were three stages in logistics development which were relied on the change of logistics vendors before e-commerce appeared. Due to the Industrial Revolution, mass production was brought to the manufacturing sector, and producing firms started managing and organizing their own storage and transport. Besides, that manufacturing and circulation were further separated led to the formation of product delivery and channels from consumers and distributors. Then, the third-party logistics firms emerged owing to rapid phase of economic globalization as well as clear social labor separation. These aspects made a contribution to the logistics outsourcing to reduce expenses and improve efficiency, according to Barreto et al. [6]. The roles of logistics have been redefined with e-commerce development: it is not a tool for building manufacturing-sale relations but is to approach customers directly. The rapid growth of express distribution enterprises based on the network along with storage and logistics companies, for example, became a significant connection between producers and final consumers [7][8].

Current tendencies and developments (such as Smart Data, Data Analytics, Big Data, and Industry 4.0) result in the popularity of data and information. However, the merit could become a disadvantage for an enterprise when too much detrimental data were provided to their setting. Information logistics aims to transfer the right element of information, in the right form, at the right place, and at the right time [9]. It is essential to handle and provide efficient information so that the aforementioned goal could be achieved, which then is significant for manufacturing. It can be defined that big data logistics is an analysis and modeling of the systems of (urban) transportation and delivery using huge data sets which are generated by mobile phone, GPS, business operations transactional data, along with human activities such as public transportation and social media. The logistics industry is changing to "information-related" services from "product-related" ones fundamentally. With advancements of technology with smart computing, requirements and demands are evolving every day. Therefore, more precise resource synthesis, as well as data distribution, could be facilitated by the real-time monitoring of vehicles [10].

There is an agreement with the concept of big data analytics for the acquisition, analysis, management, and storage of data collection on a large scale which was beyond conventional database software capabilities [11]. According to Zhong et al. [12], the characteristics of big data included rapid data flow. For example, in one day, more than 500 terabytes of new data were inserted into social media databases of Facebook; on New York Stock Exchange, about 1 terabyte of new commercial data was generated each day; and in-flight time with 30 minutes, more than 10 terabytes of data were generated by a single jet engine. The strategic importance of big data technology was not to master a great deal of information but to process meaningful data for achieving the added value of these data [11]. In distinct situations with sufficient information, a computer using a variety of machine learning algorithms could create a model of application to direct machines to conduct 'smart things' [13]. AI which is based on machine learning as well as big data is expected to penetrate to all spheres of life such as face recognition, navigation, e-shopping, and translation in the future while the latest revolutions in science and technology enable the industry of logistics and the internet to further integrate. The next wave of information technology (e.g cloud computing, big data analytics, and IoT) was anticipated to undergo mature time lasting for 5-10 years, and it would form logistics networks that had extended characteristics and full coverage [14]. Likewise, the digitization of the logistics would be noticeably enhanced, and novel approaches of labor division including sharing, crowdfunding, and crowdsourcing would be commonly in use. Besides, the economy of service and the experiential economy would continue to develop, AI technology would evolve quickly, and the logistics sector would be changed by the 'Intelligent Revolution.' Therefore, based on big data in the future would change lives, transform the formats of business and determine the trajectories of economic and social development. Moreover, it was clear that big data also positively affected the logistics industry, according to Walker [15].

In the epoch of big data, numerous data existed in the logistics sector daily, especially the entire logistics such as warehousing, distribution, packaging, transportation as well as other links where each link built a comprehensive flow of information. In case these data could not be processed in time by the logistics companies, they would face a waste of resources as well as a data disaster. Recently, many experts from both domestics and foreign countries tend to concentrate on the big data application in the industry of logistics, so adapting to the varied environment of the market and dealing with harsh market competition was essential for companies. Furthermore, other aspects of the logistics industry could be immensely affected such as operational management, strategic decision-making, service innovation, management of customer relationships, brand management, etc. Accordingly, companies could maximize the distribution of logistics services, speed up logistics sector upgrades and transformations as well as fulfill the information age requirements. With regard to big data, in the competitive advantage of the market, it was necessary that the logistics industry grab opportunities to enhance ability and efficiency, along with employ data resources.

Indeed, the logistics industry is among the best positioned to take advantage of Big Data's methodological developments and analytical capabilities. These days, transport and logistics companies are increasingly developing large and comprehensive data sets whilst handling vast flows of individuals and goods through the growing digitalization of these spheres. Millions of shipments are daily tracked and gathered in the global distribution and transportation networks to get information of the location, size, material, origin, weight, and destination. Thus, they provide precious big data sets and open up new opportunities in customer experience, operational efficiency, and novel models of business for enterprises based on data. In this article, the contents are presented in the following order: the most general overview of Big data and 4th generation logistics are presented in part 2; Part 3 is a detailed analysis of the challenges and opportunities of Big data in logistics; Part 4 clarifies potential applications of Big data in logistics 4.0.

Overview of Big Data and Logistics 4.0

Big Data

Big data (BD) is an inclusive definition for general data collection. It could be structured or unstructured, and BD is so broad and varied that it is virtually impossible to process BD using traditional systems of processing data. Thus, Big Data was defined by Provost and Fawcett as such huge sets of data that processing data was beyond the ability of traditional systems, leading to the requirement of novel technologies [16]. Big data was generated largely by computers, so it was normally shown as machine data. As regards scalability, big data was shown in four factors by IBM including (1) volume, (2) velocity, (3) variety as well as (4) veracity which was depicted in Fig. 1. The above-mentioned five "V" pillars were central elements of big data generating a brand-new "V" pillar known as a value when a large amount of meaningful information could be achieved by the organizations from big data through deeper outlooks from outstanding data analytics [17].

That definition was common due to its emphasis on the necessity and meaning of big data which covered exploring enormous hidden data values with massive size, different types, and fast generation. Altendorfer-Kaiser [18] built a 5Vs model with the addition of veracity into crucial characteristics of BD. Besides, big data was mentioned as high speed, high volume, and highly diverse information in which new processing techniques were required to discover insights, improve decision making, and optimize the process. Therefore, a set of data could also be big data in case it was capable of conduct storage, management, capture, visualization, distribution, and analysis through emerging technologies [19]. In 2014, distinct existing platforms to carry out big data analytics were discussed by Reddy and Singh who also studied various hardware platforms as well as the software framework described in the task support [20]. Whereas, in 2013, the characteristics of big data were emphasized by Ankara, Sinanc, Sagiroglu, and Turkey who then explained how important big data was in offering helpful information for enterprises [21]. Moreover, the big data overview including scope, methods, content, samples, privacy, and pros and cons was also mentioned. A literature review and framework tutorial for big data analysis platforms were presented by Han Hu, Tat-Seng, Yonggang, and Xuelong for non-expert readers in 2014 [22]. In their paper, not only the definition but also the challenges of big data were shown. Then, a systematic framework was displayed for sequentially dividing data systems into four modules including generation, acquisition, storage, and analytics of data. In 2015, changes in computing models along with the programming paradigm to address the issues and application of big data were presented by Veljko Milutinovic, Veljko Milutinovic, Anton Kos, and Jakob Salom [23].

Figure 1. Core components of Big data

In recognition of advancement, applications, and new investigations on big data, the BI&A framework was innovated [24]. BD was considered one of five core foundation techniques and new fields for analytics study, along with Web Analytics, Mobile Analytics, Text Analytics, and Network Analytics. Additionally, the term BD was used to show data exploiting as well as analyzing statistics with the use of BI&A technologies [24]. Recently, this sphere has become prominent since fast growth in gathering and creating data which normally came from the next sources of data was focused on by companies [25]. That around 2.5 Quintillion bytes of data were generated daily posed a challenge for companies, which was not about collecting and managing an immense number of data, but about the way organizations extract useful value from that vast source of data [26]. To protect intelligence from the increase in data level, organizations needed to comprehend the BD meaning and potential values, which then turned into business advantages [27], [28]. Nevertheless, companies had difficulty realizing that potential. According to the results of a survey on 720 businesses around the world, 64% of attendants planned to invest in BD projects in 1-2 years, but below 8% of respondents proposed solutions [29]. In a more recent study, to improve experience for customers, make current processes suitable, reduce expenses, and conduct more targeted marketing, the aim of BD projects was defined and related to location data analysis (70%) or freeform text (64%). However, many companies were unsure whether the return on investments was good or bad [30].

Logistics 4.0

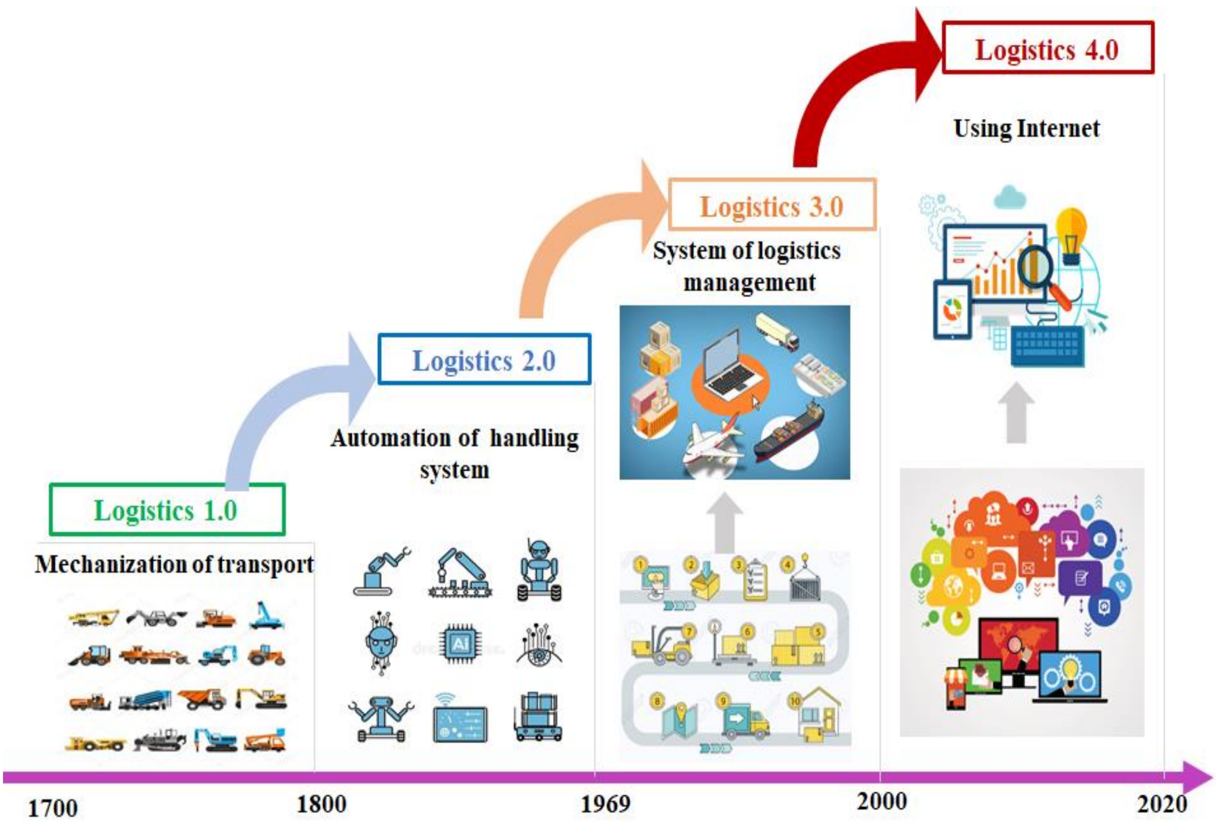

Figure 2. Evolution of Logistics (adapted from Wang, 2016)

With the shift in society, industry, and technology, logistics also changed, so due to the fourth revolution in industry and technological innovations of the 21st century, logistics 4.0 was created. The ICT development allowed novel approaches to exchange data, integrate horizontal and vertical value chains as well as form brand-new models of business. 2011 observed the first appearance of the term Logistics 4.0 which was like support and response to Industry 4.0. Other terms have also emerged these days such as Procurement 4.0, Warehousing 4.0, Supply Chain 4.0, Marketing 4.0, Order Management 4.0, Distribution 4.0, Inventory Management 4.0, and so on, which reflected the logistics sector's response to

Industry 4.0 growth and requirements. Logistics 4.0 could complement industry 4.0 processes, from market processing and production planning to smart goods distribution to final consumers. Digitalizing the activities and processes of logistics along with the use of digital logistics were considered the resolution, and according to [31], digitalization of logistic systems features included Adaptiveness, Autonomous, Cooperation, Cognition, Integration, Connectivity. The basis of logistics 4.0 was software systems and the newest ICT which offered Logistics control, goods flow realization, and information flow realization [32]. The management of logistics involved all logistical processes preparing, implementing, and controlling. The implementation of goods flows represented a series of all operations which allowed the transportation of freight flows from raw material sources through product distribution to the end-users. Finally, for supporting the two above-mentioned factors, flows of information were necessary [33].

Logistics 4.0, as its elements allowed intelligent process management, was known as smart logistics. Based on [34] logistics 4.0 components consisted of Real-time location, Connectivity and integration, analysis and processing data, Automatic identification, Business services, and Automatic data collection (Fig. 2). The automatic detection and collection of data in real-time of all items and participants in the processes of logistics enabled quality control, scheduling, and optimization. Besides, the processing and analysis of data built new insights, smart management conditions as well as novel business services. Among various technologies doing referred components, the most significant ones were illustrated in Fig. 3 [33].

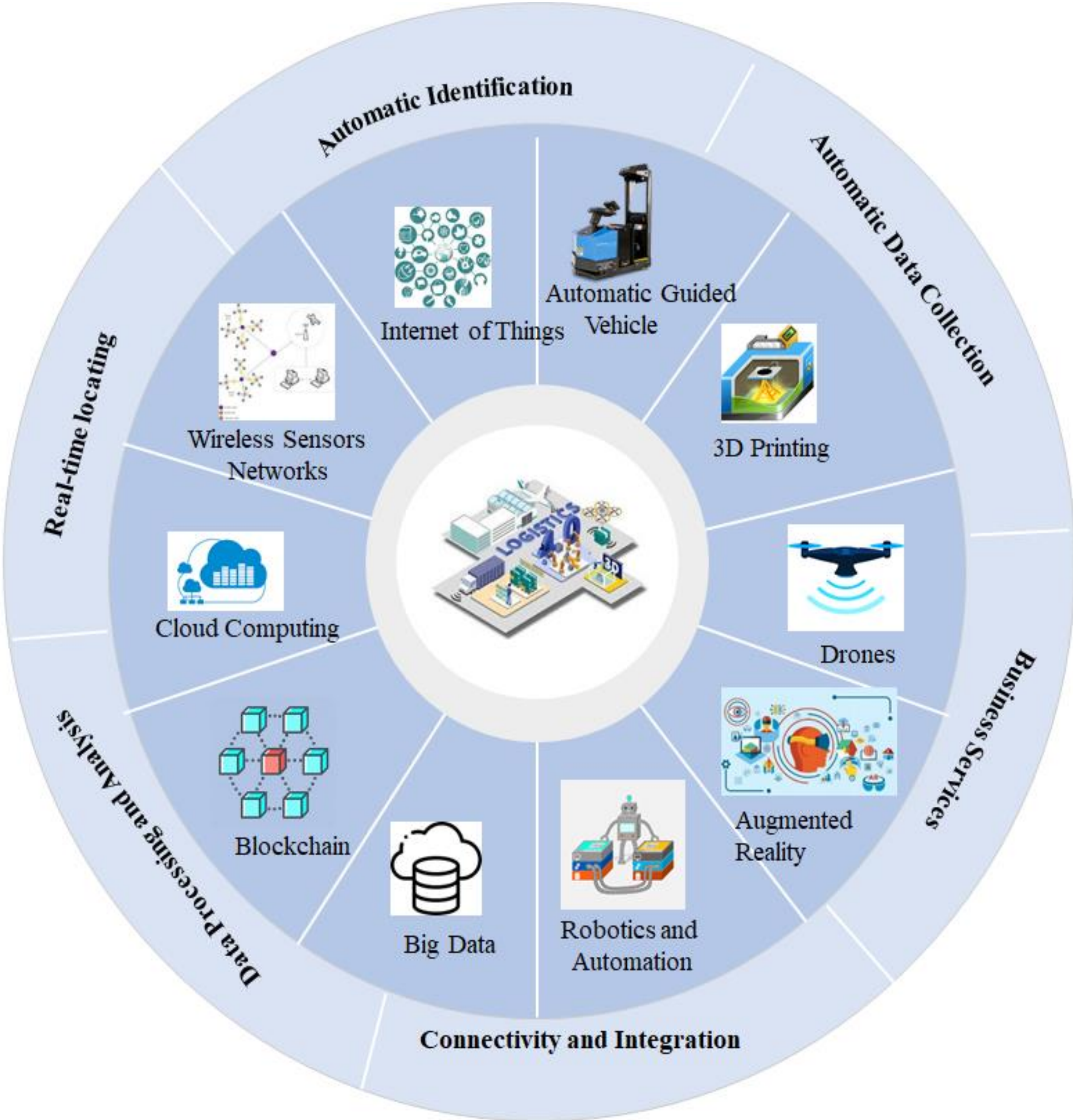

Figure 3. Components and Technologies of Logistics 4.0

Challenges and Opportunities of Big Data in Logistics

Challenges

Although big data techniques could create significant advantages in the transformation of the economy when effectively employed, there were many difficulties such as capturing, storing, sharing, searching, analyzing, and visualizing data, which must be tackled so that big data capabilities could be exploited. Among challenges, computer architecture was considered one of the biggest ones. Based on Moore's law presented by Chen and Zhang [35], CPU performance could double every 18 months. At the same pace, disk drives' performance also doubled with a small improvement in the rotational speed of disks, and there was an exponential increase in the amount of information, which greatly affected the restriction of discovering big data real-time values. Additionally, other significant difficulties were incompleteness scalability, data security, data inconsistency, and timeliness, so data needed to be created appropriately along with the applications of several pre-processing techniques (data integration, data reduction, data cleaning, and data transformation) so that noise could be alleviated and inconsistencies could be improved. Big data changed noticeably data storing and capturing such as data access mechanism, data storage device, and data storage architecture. The process of information discovery placed the usability of big data at the forefront. In this context, it was required that big data be effectively accessed and activated so that the limitations of computer infrastructure could be completely or partially broken. In storage architecture, DAS, NAS, and SAN were usually employed, but they had harsh disadvantages and restraints in distributed systems on a wide scale. To enhance the performance of data-intensive calculating, the popular way was data access optimization which contained data distribution, replication, migration, and access parallelism. When the volume of data was immense, the bottleneck in distributed networks and cloud was network bandwidth capability. Furthermore, data security was also a cloud storage-related problem. The aim of managing data was to discover and retrieve data periodically, assure the quality of data, reuse, add values and preserve data, which contained archiving, preservation, retrieval, authentication, representation, and management [19].

According to Kambatla et al. [36], inclusive analytics framework in logistics systems required the integration of customer management, advertising, supply chain management, and after-sales support. A vast quantity of multi-modal data included inventory management, video feeds based on store, customer relations, customer transactions, sales management infrastructure, advertising, customer preferences, financial data, and sentiments. Enhanced productivity could be achieved by using RFIDs to monitor inventory, supplier databases, consumer expectations, and integrated financial systems. The Big Data approach promoted the use of RFID-enabled manufacturing data in support of deciding on production logistics, according to Zhong et al. [37] as the datasets of these applications were kind of well-structured. Infrastructure and data processing were carried out under one security area, so it was easier to manage privacy and security problems. The analytics development that was able to expand the scale of multi-modal data was the main blockage in this field.

The more increasing the data volume, the more probability of confidential and valuable data; therefore, stored information for analyzing big data was susceptible to cyber-criminal [38]. Additionally, personal data available could be employed for creating values. The importance of the enormous data volume and use of big data analysis in order to create the value out of relevant data was another significant question because when such data were in use, distinct outputs could be provided to distinct categories through different pricing and quality differentiating. Compared to non – big data methods, big data analytics allowed us to determine variables with a higher correlation as well as to set prices and design services depending on these variables. It could show the combination of non-structured and structured data for diverse sources and hidden links between outer irrelevant data. According to Chen and Zhang [35], security issues covered personal privacy protection, protection of financial information, intellectual property protection, and commercial secrets. The law on data protection was introduced in many developed and developing nations. However, the problem of data security was more difficult to handle regarding big data applications due to the enormous quantity of big data and heavy security workload.

Opportunities

All forms of logistics activities included supply chain management, air, land, and maritime logistics, express delivery, freight logistics, express delivery, reverse logistics, which was enhanced through applying modern technology in transport chains. Novel technologies led to a digital shift of logistics and transport; thus, growing datasets were created. Being one of the most common tendencies in the industry of transport and logistics, the IoT benefited from high communication technologies including M2M in order to connect most objects to the Internet [39]. Besides, they could apply IoT for communication among cars as well as various applications of Intelligent Transport Systems. Especially, vehicles and other objects could be more connected thanks to the increasing use of sensors, GPS, RFID, and WIFI; therefore, logistics and transport played an essential part in the sources of data sets. In fact, in terms of resources and management of capabilities, these data could pose an unavoidable challenge; nevertheless, from the data potential, great opportunities could be achieved to form valuable business models and to enhance both operational efficiency and customer experience [40].

Indeed, data-driven companies on transport and logistics could benefit from numerous data on delivery, pickup, and transport chain to expand services and create new assets of information and business models. The geographical coverage of the continuously ongoing distributed fleet which was connected with mobiles, cameras, or sensors was a useful source to provide the new customers with rich datasets and information. In different cases of environmental intelligence use, the popularity of this huge fleet of transport and delivery could be utilized [41]. Logistics providers could provide environment agencies, real-estate developers, and authorities with measurement data (i.e. noise, temperature, traffic density, humidity, and pollution) as they created helpful environmental statistics for environmental supervising activities and city planning. The acquisition of customer information was another noticeable factor of BD analytics. In fact, valuable data collected from the network of distribution and transport enabled to manage the customer relations and understand their demands. Combining various sources of data could gain detailed information of customers, which enabled transport firms and logistics providers to comprehensively estimate customer satisfaction as well as minimize customer loss [42].

Collecting customer details played an important role in big data analytics and distribution network data provided significant value for customer relation research and management. The applications of big data methods helped comprehend the demand and satisfaction of customers; thus, logistics providers had to combine multiple data to attain the comprehensive information of customers. It was essential that big data analytics offer an inclusive insight of operational performance and customer interactions as well as ensure sender and receiver's satisfaction. With the aim of attaining precise findings from an assessment of customer feedback, logistics providers needed to get details from various touchpoints. According to Robak et al. [43], the issue of further study in supply chain management and logistics systems could be investigated from the aspect of distinct groups of stakeholders and elements of managerial business, in which the major functions of business included transportation management, forecasting, human resources, inventory management, and transport. Thanks to Big Data, many problems (responding to customer experience in time, inventory, time delivery forecast, managing relationships between suppliers and customers, real-time planning of ability) could be tackled. Based on the study by Rozados and Tjahjono [44], it was important that companies in supply chain management consider data as strategic assets, not just information assets because they were able to know economic values by the use of Big Data analytics through activities of producing revenue.

Potential Applications of Big Data in Logistics 4.0

All forms of logistics activities included supply chain management, air, land, and maritime logistics, express delivery, freight logistics, express delivery, reverse logistics, which was enhanced through applying modern technology in transport chains. Novel technologies led to a digital shift of logistics and transport; thus, growing datasets were created. Being one of the most common tendencies in the industry of transport and logistics, the IoT benefited from high communication technologies including M2M in order to connect most objects to the Internet [39]. Besides, they could apply IoT for communication among cars as well as various applications of Intelligent Transport Systems. Especially, vehicles and other objects could be more connected thanks to the increasing use of sensors, GPS, RFID, and WIFI; therefore, logistics and transport played an essential part in the sources of data sets. In fact, in terms of resources and management of capabilities, these data could pose an unavoidable challenge; nevertheless, from the data potential, great opportunities could be achieved to form valuable business models and to enhance both operational efficiency and customer experience [40].

The Council of Supply Chain Management Professionals [45] noted that a massive data amount was produced by global logistics when carriers, logistics providers, and shippers handled logistics activities. For example, big data generated from EDI transactions, mobile devices, and RFID tags [46] could be exploited for the purpose of logistics planning. This was related to goods distribution from supply points such as warehouses and manufacturing places to demand points including retailer sites via intermediate points of storage like distribution centers. Problems in the logistics planning could emerge as issues in network flows, in which each arc was a transport mode with a particular capacity and duration [47-49]. In distribution networks, logistical data was created from various sources including supply capacity prediction for suppliers' plants, network capacity, shipping expenses, and demand forecasts of demand points [50]. In order to build the flexibility of the supply chain in logistics activities, it was crucial to use predictive analytics tools due to demand uncertainty and supply disruptions [51].

Optimization of equipment and crew routing was essential in logistics planning. The issue of vehicle routing seeks to the sequence optimization of accessed nodes (i.e. for a returns truck, for parcel delivery vehicles, or both) [52], [53]. The optimum order took into account distances between each node pair, left turns, the anticipated traffic volume, and other limitations imposed on routes including delivery and pickup times [54]. Nevertheless, in the international logistics network, it was complex for transport planning and distribution activities due to vehicle capacities, time of delivery and pickup, multiple vehicles, and tour-length limitations [55]. The use of analytics techniques and methods could optimize the routing of products, crew, and vehicles [56-58] so that the transport expenses and margins could be balanced, and safety and maintenance could be noticed.

The combination between cloud manufacturing and IoT was focused to attain ultimate intelligent manufacturing. Numerous data would be created thanks to IoT technologies like RFID employed in factories. These data were nuanced, abstract, and changeable, so data containing a variety of helpful information was difficult to fully utilize. In this study [59], a visualization method was presented for the Big Data from Cloud Manufacturing RFID-enabled shopfloor logistics. Because of the manufacturing logic as well as time series, an advanced RFID-Cuboid was in use to reconstruct the RFID raw data. There were some emphasized contributions. First, a practical method for integration of Cloud Manufacturing and IoT was launched so that the traditional industry could be upgraded and transformed for a smart future. Second, through the use of timestamps and production logic, a model of RFID-Cuboid was presented for chaining the RFID data, leading to the interpretation of data. Third, the viability and practicality of the mentioned visualization method were recorded in a real-life scenario, which made it easy for various end-users to do their operations every day. From this scenario, we could learn meaningful lessons to implement big data analytics and IoT-supported Cloud Manufacturing in the industrial sphere.

It was important to make a good decision based on BI so that competitiveness would be ensured for sustainable development. That communication and information technology rapidly developed increased the importance of bog data analysis and collection, which led to a significant rise in BDA. Nevertheless, not much research was related to BI because enterprises hardly comprehend and use terms integrally; thus, the study aim [60] consisted of two parts. Firstly, we studied documents on big data, BDA, and BI to know that these approaches were not separated but an integrated system of supporting the decision. Secondly, the way enterprises utilized BDA and big data, in reality, was explored, along with the use of BI through an investigation of logistics classifying and processing of a common carrier company. We concentrated on the expense efficiency related to data analysis or simulation, data collection as well as the findings from real-life applications. Based on this paper's results, firms could attain management efficiency by the use of big data through effective BI instead of making any investment in other equipment. Additionally, companies could gain indirect experience which allowed them to reduce error and trial and remain or raise competitiveness.

Some developments were expected in the maritime industry for efficient management of operational processes with the introduction of technologies supporting Industry 4.0. In the maritime logistics chain, seaports were the most essential point due to their complicated and multimodal feature, so port-related parties needed to communicate for operation improvement. As main agents of digital seaports, EDI and PCS have recently shown shortcomings in exchanging information efficiently, timely, securely, and accurately, so this led to high operating expenses, low performance, and poor management of the resource. Due to the aforementioned reasons, this paper showed the SDS based on the model of IDS architecture model so that a space to share data securely could be enabled and a multimodal terminal of intelligent transport could be boosted. The IDS connector was implemented by each port-related party to join SDS as well as share the data. An architecture of big data was integrated on SDS for enormous data management that was exchanged in the SDS and helpful information extraction to enhance decision making. This architecture was assessed by permitting container terminals and port authorities to interchange data with shipping firms, which led to the development of KPIs through employing functionalities of big data architecture. The dashboard displayed KPIs so that results could easily be explained to plan ship operations. The communication among relevant parties might be enhanced by the SDS settings through the reduction in transaction expenses, improvement of information quality, and effectiveness [61].

For the global economy, worldwide seaports are significant. International container traffic has increased by 10 percent on average yearly since 1990. Similarly, the main issues related to logistics and techniques emerged widely due to the great increase of vessel sizes; thus, maritime and shipping logistics could take advantage of big data and other new digital technologies. However, besides positive impacts (such as energy-saving, safety, and efficiency), digitization in logistics and maritime also posed some risks such as cybercrime and data abuse. According to a systematic review of the document, in this paper an overview on the recent status of digitization in maritime logistics was offered, emerging problems were discussed and enhancement potential was shown. Thus, grabbing the potential of development was of importance to exploit benefits. Nevertheless, it was still at the beginning of the investigation, so there existed a lack of practical and theoretical research, along with explanatory methods for restructuring and action recommendation [62].

The Boost 4.0 project currently was launched for designing and implementing Big Data middleware to support IDS with the main motivation of narrowing down the gap between the management of big data and IDS architecture [63], [64]. The results of the project were expected to reveal by the end of 2021. Besides, in existing documents on the maritime industry, hardly any big data architecture was presented [65], [66]. Using Lambda processing architecture was the key method of these architectures since it was effective in fulfilling the requirements of high availability, efficiency, and scalability [67]. Nevertheless, IoT requirements for the management of big data or the employment of big data lifetime models for the designs were not reviewed in those architectures. Recommendations were presented by the ITU [68] to be noticed in designing the architecture of big data for IoT. In addition, Demchenko et al. [69] proposed the BDLM model that offered important advantages for reusability of data at any phase of life cycle and for immense data decline at an early period.

Therefore, big data architecture was a basis for sorting related information from exchanged data with the aim of enhance the operations of seaports [61].

The purpose of this research [70] was to recognize the present state of BDA on various structural levels of SCM in Brazilian companies. In particular, this paper concentrated on knowing awareness of BDA in Brazilian enterprises and released a framework for maturity analysis of companies in conducting BDA projects in SCM or logistics. A questionnaire survey was carried out on SCM levels of one thousand companies, in which 155 out of 272 reports were valid, accounting for 15.5%. Accordingly, it could determine the acquaintance of Brazilian companies on BDA, obstacles when applying BDA projects, and relations between BDA knowledge and supply chain levels. Thus, a framework was presented to adopt BDA schemes in SCM. External value because of constraints in the generalization of findings from a Brazilian perspective from experimental sampling was not covered in this research. Further investigations should understand this domain and concentrate on how big data affected networks or supply chains in emerging areas like Latin America. Moreover, in this study, insights for developing activities relating to SCM and big data were provided and consistent, functional guidelines were displayed via triangle framework of BDA-SCM as an extra method in conducting BDA projects from SCM perspective [70].

The major aims of big data analysis were divided into six groups (Fig. 4) by the Gartner Group via the value model of big data (2015) [71]. Information on product process efficiency, operational excellence, customer insight, digital marketing, and digital products service was included in this model which reviewed main instances of big data employed by domestic and foreign firms through reusing major aims of big data analysis as enhancing internal processes, creating novel value positions, managing customer relationships and boosting efficiency [71]. Big data was an approach extracting data value as well as analyzing the findings and it contained a wide range of non-structured or structured data beyond the capacity of current database management software. There were numerous cases in logistics in which big data technologies were in use. Amazon, for instance, employed big data technology to manage inventory systems and predict delivery services. As a result, Amazon predicted accurately the changes in buying patterns of customers in real-time as well as apply these in the policy of inventory. Thanks to the application of big data technology in managing inventory systems, inventory quantity was effectively handled, efficiency was maximized, costs were declined and profits were created. Furthermore, the old order lists of buyers and attractive goods in shopping carts were also analyzed because this information allowed consumers to get items early after the order by a delivery forecast from the nearest storage even when the purchase was still uncertain. ZARA, a fashion brand, applied RFID tags for all clothes to determine the state of the stock, garments often worn by buyers, and preferences of consumers. Then, these preference data were collected via SNS and online stores before sent to the data center which eventually analyzed enormous data and utilized it for generating new products. Additionally, the cooperation between ZARA and MIT allowed reducing unneeded inventory. This was attained by the development of an inventory distribution system based on big data in which they could analyze data on inventory and sales from shops worldwide in real-time. Based on big data, ZARA decided to the analysis of demand prediction by items, sales tendencies by the shop, and the relation between the shown number of products and volume. Thanks to the accuracy of deciding with the use of data, a non-stock operation policy was realized [71], [72]. In 2011, a resolution for Supply Chain and Logistics was developed by Samsung SDS, known as "Cello" which was based on IT system establishment and experience in integrated logistics [72].

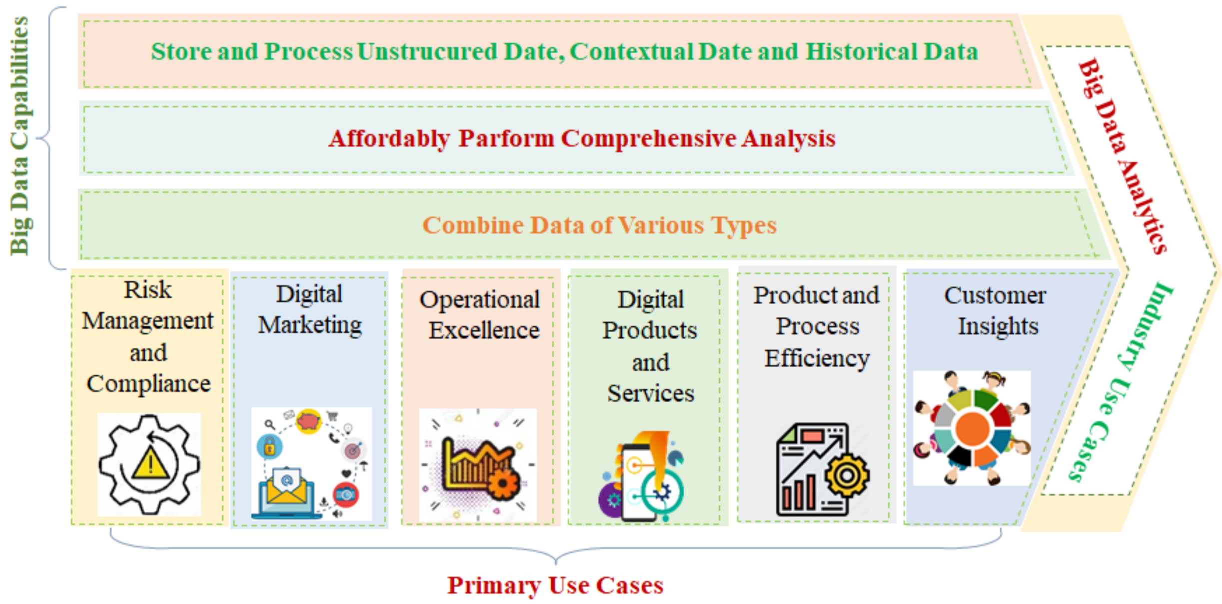

Figure 4. Big data analytics and industry use case

In this article [73], current distribution modes applied by e-commerce businesses in China were analyzed. This research compared and examined various logistics distributor modes that e-commerce companies had to deal with (such as new big data features, challenges, and advantages) according to the experimental investigation of an electronic shopping center at JD.com (or Jing-Dong). In order to examine the distribution choice mode of e-commerce businesses, the AHP approach, as well as entropy values, were adopted, and the TOPSIS method was employed for verifying models. Our findings in this study analysis brought deep managerial knowledge to practitioners whose job was related to e-commerce logistics distribution [73].

Our study explored the support by systems of big data that were developed in the cloud to authorize emerging logistics services via improving synergies among 3/4PL so that sustainable and highly integrated logistics supply chain services could be founded. Nevertheless, instead of handling the data accuracy and quality, the application of big data had restricted abilities to offer effective logistics services. The key result of the article was defining an architectural framework along with related study and development programs for cloud computing application to develop and utilize BDLBP in services of managing supply chain network. The abilities in the BDLBP played an important role in the logistics network and stakeholders. Real-time route optimization and operational capacity planning were two in three key capabilities, and they were formed according to operational studies as well as expanded to the scale of uncertain cases. Regarding the third ability, logistics network planning was investigated to include this ability supported by the analysis of big data in the cloud [74].

It is a new tendency to apply big data technologies in logistics and transport applications, but the fast growth of IT, Robotics, Data Science, and AI leads to the rapid appearance of novel definitions, creating opportunities for new prospects in the future. Moreover, Industry 4.0 is considered a scheme for the future that integrates IT, IoT, AI technologies in production, resulting in new terms emerging which includes the Smart Factory and Smart Technologies. The new abilities of self-governing process and operation management allowed independent decisions on the process to develop in the transport and logistics field, so-called Smart Logistics. In Smart Logistics, modern smart and self-directing vehicles in traffic infrastructure based on IoT and BD were also a primary tendency that could offer automatic and flexible logistics resolutions. The “Future Truck 2025”, for instance, was designed by the brand Mercedes was a self-driving truck that could change the shipping industry's future. In this background, logistics and transport, and big data went together because the widely shared information on vehicles’ sensor data, traffic, and the weather could lead to more effective flows of logistics and transport.

The Anticipatory Logistics, based on predictive analysis of big data, enabled logistics companies to promote the efficiency of process and quality of service through the demand prediction before requests, as well as orders, were placed so that delivery time became shorter. As for the Smart City Planning, anticipatory algorithms were adopted for matching necessary logistics resources level with the demand while with Anticipatory Shipping, logistics businesses could transport goods to distribution center which was closer to potential customers depending on the analysis of their buying behaviors, so delivery could occur with several days or even one hour. Amazon, for example, developed an Anticipatory Shipping application and it was patented as a system as well as an approach for anticipatory package shipping in 2013 [75].

Conclusion

Logistics plays a key role in boosting a country's economy. The past years have marked an extremely important transformation for the transport industry and Logistics, with many advanced technology solutions being launched. In the context of industrial revolution 4.0, technology is vital to a logistics enterprise. IoT technology is one of the technologies that increasingly influences the development of Logistic. The increasing development of technology and the application of information technology in the logistics industry have made it possible for us to collect a huge amount of data from many different sources. Big Data helps optimize capacity and enhance user experience to minimize risky activities and create a perfect business model. Besides, Big data helps with resource usage and quality of the process, performance to increase speed, and transparency in decision-making helps to improve the efficiency of business operations. Using Big Data in supply chain management can help businesses more accurately forecast demand, better understand customer buying cycles, and calculate future stock output based on old data. Data cannot be seen only as information sitting on the cloud but must be seen as a special asset to the organization. Thus, it can be seen that technology trend 4.0 will be the core foundation for logistics to take off in the future, and not only big companies involved in solving logistics problems, but also startups will also offer breakthrough solutions in each stage of the supply chain and logistics.

Acknowledgment: This research is supported by the Academy of Politics Region II. Additionally, the authors are grateful to Ho Chi Minh City University of Technology (HUTECH) for providing the necessary facilities during the research.

Funding Statement: The author received no specific funding for this study.

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study.

References

1. S. Jafarzadeh and I. Schjølberg, “Operational profiles of ships in Norwegian waters: An activity-based approach to assess the benefits of hybrid and electric propulsion,” *Transportation Research Part D: Transport and Environment*, 65, 500–523, 2018.

2. L. Po, N. Bikakis, F. Desimoni and G. Papastefanatos, “Linked Data Visualization: Techniques, Tools, and Big Data,” *Synthesis Lectures on Semantic Web: Theory and Technology*, vol. 10, no. 1, pp. 1–157, 2020.
3. C. Qi, “Big data management in the mining industry,” *International Journal of Minerals, Metallurgy and Materials*, vol. 27, no. 2, pp. 131–139, 2020.
4. J. Fernie and L. Sparks, “Retail logistics: changes and challenges,” in *Logistics and retail management: emerging issues and new challenges in the retail supply chain*, 3rd ed, Kogan page publishers, pp. 3-37, 2018.
5. Z. Wang, M. Xue, Y. Wang, M. Song, S. Li *et al.*, “Big data: New tend to sustainable consumption research,” *Journal of Cleaner Production*, vol. 236, p. 117499, 2019.
6. L. Barreto, A. Amaral and T. Pereira, “Industry 4.0 implications in logistics: an overview,” *Procedia Manufacturing*, vol. 13, pp. 1245–1252, 2017.
7. Kawa, “Fulfillment service in e-commerce logistics,” *Scientific Journal of Logistics*, vol. 13, no. 4, pp. 429-438, 2017.
8. Z. Yan, H. Ismail, L. Chen, X. Zhao and L. Wang, “The application of big data analytics in optimizing logistics: a developmental perspective review,” *Journal of Data, Information and Management*, vol. 1, no. 1, pp. 33–43, 2019.
9. R.Z. Khayrullin, “The system for statistical analysis of logistic information,” in *Proc. Vestnik MGSU*, vol. 5, pp. 133-140, 2015.
10. N. Hoang Phuong, N. Minh Tuan and P. Minh Tuan, “Logistics Revolution for e-commerce in Vietnam : A Brief Review,” *International Journal of e-Navigation and Maritime Economy*, vol. 13, pp. 50–62, 2019.
11. Y. Wang, L. Feng, H. Chang and M. Wu, “Research on the impact of big data on logistics,” in *Proc. MATEC Web of Conferences*, vol. 100, p. 2015, 2017.
12. R.Y. Zhong, S.T. Newman, G.Q. Huang and S. Lan, “Big Data for supply chain management in the service and manufacturing sectors: Challenges, opportunities, and future perspectives,” *Computers & Industrial Engineering*, vol. 101, pp. 572–591, 2016.
13. M. Kataria and M. P. Mittal, “Big data: a review,” *International Journal of Computer Science and Mobile Computing*, vol. 3, no. 7, pp. 106–110, 2014.

14. X. Li, “Development of intelligent logistics in china,” in *Contemporary logistics in China*, Springer, 2018, pp. 181–204.

15. M. Viktor and K. Cukier, “Big Data: A Revolution That Will Transform How We Live, Work, and Think,” *American Journal of Epidemiology*, 2014.

16. F. Provost and T. Fawcett, “Data science and its relationship to big data and data-driven decision making,” *Big data*, vol. 1, no. 1, pp. 51–59, 2013.

17. S. Altendorfer-Kaiser, “The influence of big data on production and logistics,” in *Proc. IFIP International Conference on Advances in Production Management Systems*, 2017, pp. 221–227.

18. A.E. Hassanien, A.T. Azar, V. Snasael, J. Kacprzyk and J.H. Abawajy, “Big data in complex systems,” in *SBD*, vol. 9, Springer, 2015.

19. B. Mikavicaa, A. Kostić-Ljubisavljevića and V. Radonjić, “Big data: challenges and opportunities in logistics systems,” in *Proc. of the 2nd Logistics International Conference*, 185–190, 2015.

20. Desfiandi, A., Suman Rajest, S., S. Venkateswaran, P., Palani Kumar, M., & Singh, S. (2019). Company Credibility: A Tool To Trigger Positive Csr Image In The Cause-Brand Alliance Context In Indonesia. Humanities & Social Sciences Reviews, 7(6), 320-331.

21. H. Zhou, G. Sun, S. Fu, X. Fan, W. Jiang *et al.* “A distributed approach of big data mining for financial fraud detection in a supply chain,” *Computers,Materials & Continua*, vol. 64, no. 2, pp. 1091-1105, 2020.

22. H. Hu, Y. Wen, T.-S. Chua and X. Li, “Toward scalable systems for big data analytics: A technology tutorial,” *IEEE access*, vol. 2, pp. 652–687, 2014.

23. N. Trifunovic, V. Milutinovic, J. Salom and A. Kos, “Paradigm shift in big data supercomputing: dataflow vs. controlflow,” *Journal of Big Data*, 2, 1, 1–9, 2015.

24. H. Chen, R.H.L. Chiang and V.C. Storey, “Business intelligence and analytics: From big data to big impact,” *MIS quarterly*, pp. 1165–1188, 2012.

25. F.J. Alexander, A. Hoisie and A. Szalay, “Big data,” *IEEE Annals of the History of Computing*, vol. 13, no. 06, pp. 10–13, 2011.

26. H. Zhou, G. Sun, S. Fu, W. Jiang and J. Xue, "A scalable approach for fraud detection in online e-commerce transactions with big data analytics," *Computers, Materials & Continua*, vol. *60*, no. 1, pp. 179-192, 2019.

27. M.M. Gobble, "Big data: The next big thing in innovation," *Research-technology management*, vol. 56, no. 1, pp. 64–67, 2013.

28. McAfee, E. Brynjolfsson, T.H. Davenport, D.J. Patil and D. Barton, "Big data: the management revolution," *Harvard business review*, vol. 90, no. 10, pp. 60–68, 2012.

29. L. Kart, N. Heudecker and F. Buytendijk, "Survey analysis: big data adoption in 2013 shows substance behind the hype," *Gartner Report GG0255160*, vol. 13, 2013.

30. J. Hopkins and P. Hawking, "Big Data Analytics and IoT in logistics: a case study," *The International Journal of Logistics Management*, 2018.

31. Y. Kayikci, "Sustainability impact of digitization in logistics," *Procedia manufacturing*, 21, 782–789, 2018.

32. Z. Zhang, J. Chen, L. Chen, Q. Liu, L. Yang *et al.* "A scalable method of maintaining order statistics for big data stream," *Computers, Materials & Continua*, vol.60, no.1, pp. 117-132, 2019.

33. G. Radivojević and L. Milosavljević, "The concept of logistics 4.0," in *Proc. 4th Logistics International Conference*, pp. 23–25, 2019.

34. K. Wang, "Logistics 4.0 solution-new challenges and opportunities," in *Proc. 6th International Workshop of Advanced Manufacturing and Automation*, pp. 68–74, 2016.

35. C.L.P. Chen and C.Y. Zhang, "Data-intensive applications, challenges, techniques and technologies: A survey on Big Data," *Information sciences*, 275, 314–347, 2014.

36. K. Kambatla, G. Kollias, V. Kumar, and A. Grama, "Trends in big data analytics," *Journal of parallel and distributed computing*, vol. 74, no. 7, pp. 2561–2573, 2014.

37. R.Y. Zhong, G.Q. Huang, S. Lan, Q.Y. Dai, X. Chen *et al.*, "A big data approach for logistics trajectory discovery from RFID-enabled production data," *International Journal of Production Economics*, vol. 165, pp. 260–272, 2015.

38. N. Kshetri, "Big data′ s impact on privacy, security and consumer welfare," *Telecommunications Policy*, vol. 38, no. 11, pp. 1134–1145, 2014.

39. W. Liu, Y. Tang, F. Yang, C. Zhang, D. Cao and G. Kim, "Internet of Things based solutions for transport network vulnerability assessment in intelligent transportation systems," *Computers, Materials & Continua*, vol. 65, no. 3, pp. 2511-2527, 2020.
40. T. Borgi, N. Zoghlami and M. Abed, "Big data for transport and logistics: A review," in *Proc. IC_ASET*, pp. 44–49, 2017.
41. S. Zhou, W. Liang, J. Li and J. Kim, "Improved VGG model for road traffic sign recognition," *Computers, Materials & Continua*, vol. 57, no. 1, pp. 11-24, 2018.
42. K. Coussement, S. Lessmann and G. Verstraeten, "A comparative analysis of data preparation algorithms for customer churn prediction: A case study in the telecommunication industry," *Decision Support Systems*, vol. 95, pp. 27–36, 2017.
43. S. Robak, B. Franczyk and M. Robak, "Research Problems Associated with Big Data Utilization in Logistics and Supply Chains Design and Management," in *Proc. FedCSIS*, 245–249, 2014.
44. I.V. Rozados and B. Tjahjono, "Big data analytics in supply chain management: Trends and related research," in *Proc. 6th International Conference on Operations and Supply Chain Management*, pp. 10–13, 2014.
45. H. Sun and S. McIntosh, "Analyzing cross-domain transportation big data of New York City with semi-supervised and active learning," *Computers, Materials & Continua*, vol. 57, no. 1, pp. 1-9, 2018.
46. S. Swaminathan, "The effects of big data on the logistics industry," *Profit Oracle*, 2012.
47. M. Dong and F.F. Chen, "Performance modeling and analysis of integrated logistic chains: An analytic framework," *European Journal of Operational Research*, 162, 1, 83–98, 2005.
48. S. Jharkharia and R. Shankar, "Selection of logistics service provider: An analytic network process (ANP) approach," *Omega*, vol. 35, no. 3, pp. 274–289, 2007.
49. C.S. Grewal, K.K. Sareen and S. Gill, "A multicriteria logistics-outsourcing decision making using the analytic hierarchy process," *International Journal of Services Technology and Management*, vol. 9, no. 1, pp. 1–13, 2008.

50. M. Najafi, K. Eshghi and W. Dullaert, “A multi-objective robust optimization model for logistics planning in the earthquake response phase,” *Transportation Research Part E: Logistics and Transportation Review*, vol. 49, no. 1, pp. 217–249, 2013.

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

52. L. Özdamar and O. Demir, “A hierarchical clustering and routing procedure for large scale disaster relief logistics planning,” *Transportation Research Part E: Logistics and Transportation Review*, vol. 48, no. 3, pp. 591–602, 2012.

53. M. Drexl, “Applications of the vehicle routing problem with trailers and transshipments,” *European Journal of Operational Research*, vol. 227, no. 2, pp. 275–283, 2013.

54. T. Vidal, T.G. Crainic, M. Gendreau and C. Prins, “Heuristics for multi-attribute vehicle routing problems: A survey and synthesis,” *European Journal of Operational Research*, vol. 231, no. 1, pp. 1–21, 2013.

55. X. Li, P. Tian and S.C.H. Leung, “Vehicle routing problems with time windows and stochastic travel and service times: Models and algorithm,” *International Journal of Production Economics*, vol. 125, no. 1, pp. 137–145, 2010.

56. C. Novoa and R. Storer, “An approximate dynamic programming approach for the vehicle routing problem with stochastic demands,” *European journal of operational research*, vol. 196, no. 2, pp. 509–515, 2009.

57. H. Lei, G. Laporte and B. Guo, “The capacitated vehicle routing problem with stochastic demands and time windows,” *Computers & Operations Research*, vol. 38, no. 12, pp. 1775–1783, 2011.

58. Minis and A. Tatarakis, “Stochastic single vehicle routing problem with delivery and pick up and a predefined customer sequence,” *European journal of operational research*, vol. 213, no. 1, pp. 37–51, 2011.

59. R.Y. Zhong, S. Lan, C. Xu, Q. Dai and G.Q. Huang, “Visualization of RFID-enabled shopfloor logistics Big Data in Cloud Manufacturing,” *The International Journal of Advanced Manufacturing Technology*, vol. 84, no. 1–4, pp. 5–16, 2016.

60. D.H. Jin and H.J. Kim, “Integrated understanding of big data, big data analysis, and business intelligence: a case study of logistics,” *Sustainability*, vol. 10, no. 10, pp. 3778, 2018.

61. D. Sarabia-Jácome, C.E. Palau, M. Esteve and F. Boronat, “Seaport data space for improving logistic maritime operations,” *IEEE Access*, vol. 8, pp. 4372–4382, 2019.

62. M. Fruth and F. Teuteberg, “Digitization in maritime logistics—What is there and what is missing?,” *Cogent Business & Management*, vol. 4, no. 1, p. 1411066, 2017.

63. R.M. Holom, K. Rafetseder, S. Kritzinger and H. Sehrschön, “Metadata management in a big data infrastructure,” *Procedia Manufacturing*, vol. 42, pp. 375–382, 2020.

64. S. Cho, G. May, I. Tourkogiorgis, R. Perez, O. Lazaro *et al.*, “A hybrid machine learning approach for predictive maintenance in smart factories of the future,” in *Proc. IFIP International Conference on Advances in Production Management Systems*, pp. 311–317, 2018.

65. G. Vouros, C. Doulkeridis, G. Santipantakis, A. Vlachou, N. Pelekis *et al.*, “Big data analytics for time critical maritime and aerial mobility forecasting,” *Advances in Database Technology-EDBT*, vol. 2018, pp. 612–623, 2018.

66. H. Wang, O.L. Osen, G. Li, W. Li, H.N. Dai *et al.*, “Big data and industrial Internet of Things for the maritime industry in Northwestern Norway,” in *Proc. TENCON*, vol. 2016, 2016.

67. N. Marz, “Principles and Best Practices of Scalable Realtime Data Systems,” *James Warren*, vol. 328, 2015.

68. Y. Mehmood, F. Ahmad, I. Yaqoob, A. Adnane, M. Imran *et al.*, “Internet-of-things-based smart cities: Recent advances and challenges,” *IEEE Communications Magazine*, 55(9), 16–24, 2017.

69. Y. Demchenko, C. De Laat and P. Membrey, “Defining architecture components of the Big Data Ecosystem,” in *Proc. 2014 CTS*, pp. 104–112, 2014.

70. M.M. Queiroz and R. Telles, “Big data analytics in supply chain and logistics: An empirical approach,” *The International Journal of Logistics Management*, vol. 29, no. 2, pp. 767-783, 2018.

71. S.W. Kim, “Value Model for Applications of Big Data Analytics in Logistics,” *Journal of Digital Convergence*, vol. 15, no. 9, pp. 167–178, 2017.

72. D.J. Park and W.S. Kim, “Improvement of the Parallel Importation Logistics Process Using Big Data,” *Journal of information and communication convergence engineering*, 17(4), 267–273, 2019.

73. K. Zheng, Z. Zhang and B. Song, “E-commerce logistics distribution mode in big-data context: A case analysis of JD. COM,” *Industrial Marketing Management*, 86, 154–162, 2020.

74. Neaga, S. Liu, L. Xu, H. Chen and Y. Hao, “Cloud enabled big data business platform for logistics services: A research and development agenda,” in *Proc. International Conference on Decision Support System Technology*, 22–33, 2015.

75. Angeleanu, “New technology trends and their transformative impact on logistics and supply chain processes,” *International Journal of Economic Practices and Theories*, 5(5), 413–419, 2015.