# Application of Big Data in Logistics

**Conference Paper** · September 2020

DOI: 10.1145/3410352.3410785



**3 authors**, including:

Aiman Moldagulova
International University of Information Technology
**9** PUBLICATIONS **145** CITATIONS

SEE PROFILE

Ryskhan Satybaldiyeva
International University of Information Technology
**14** PUBLICATIONS **68** CITATIONS

SEE PROFILE

# APPLICATION OF BIG DATA IN LOGISTICS

Aiman Moldagulova
International Information Technology University
St. Manas 34/1
050040 Almaty
+7(701)7279025
a.moldagulova@edu.iitu.kz

Ryskhan Satybaldiyeva
International Information Technology University
St. Manas 34/1
050040 Almaty
+7(701)5509145
r.satybaldiyeva@edu.iitu.kz

Abu Kuandykov
International Information Technology University
St. Manas 34/1
050040 Almaty
+7(701)7662024
a.kuandykov@edu.iitu.kz

## ABSTRACT
This article discusses the prospect of using big data to optimize logistic processes. An analytical model for the collection and analysis of data from various sources was built. Firstly, the evolution and features of both logistics and big data were have been analyzed using the systematic review method. This was followed by discussions about the implementation of big data in logistics and the results of optimization. The article summarizes the main effects of the use of Big Data in logistics such as informatization; efficiency; quality of service; and promoting technical modernization.

## Keywords
Logistics; Big Data, Data sources, Analytical model

## 1. INTRODUCTION
Modern supply chains of companies are complex and multidimensional. They require the analysis of large amounts of data to efficiently use resources, empower and manage supply chain risks. The process of chaotic collection of operational data and spontaneous reporting is replaced by the use of analytical models aimed at optimizing logistics processes and routes, reducing costs, reducing errors from the use of manual labor, streamlining production functions and ensuring transparency of the entire supply chain.

The global leading logistics companies closely monitor all current trends in the field of IT for their application in their business. The large-scale implementation of digital technologies, such as Big Data, must be ensured in Kazakhstan. Today, several transcontinental corridors pass through the country. In general, cargo transit through Kazakhstan is growing annually. Big data technology will help to track the movement of goods in real time mode and their smooth transit, simplify customs operations, as well as organize the interaction of all parts of the logistics. Using Big data will provide carriers with high-quality analytics, identify growth reserves and reduce excess costs.

## 2. RELATED WORK
Digital technologies have allowed to accumulate huge amounts of information that can be used to make management decisions.

*ICEMIS'20 (DTESI'20),* September 14–16, 2020, Almaty, Kazakhstan

ACM ISBN 978-1-4503-7736-2/20/06$15.00

DOI: https://doi.org/10.1145/3410352.3410785

Professional third-party logistics companies have appeared. All these factors contributed to the outsourcing of logistics by producers and sellers, resulting in lower costs and increased efficiency [1], [2]). With the development of e-commerce, the logistics function has been redefined: logistics is no longer just a tool used to build links between production and sales but is now used for direct communication with consumers. For example, the rapid development of express delivery network companies and warehousing and logistics enterprises has become an important link between manufacturers and end users [3].

The latest scientific and technological advances continue to foster the deep integration of the Internet and the logistics industry. The next generation of information technologies, such as big data analytics, cloud computing and the Internet of things, is expected to enter a mature period over the next five to ten years, and full-coverage logistics networks will be formed [4]) . Digitization of logistics will be significantly improved and new methods of division of labor, such as crowdsourcing, crowdfunding and exchange, will be widely used. The economics of services and the economics of experience will deepen even further, artificial intelligence technologies will evolve rapidly, and the “intellectual revolution” will change the logistics industry. Thus, in the future, big data will more quickly transform business formats, change lives and guide the paths of socio-economic development, which will inevitably have a more positive impact on the logistics industry [5], [6],[18]-[22].

A prerequisite for producing large-scale logistics data is that several components, such as logistics elements, logistics tools, logistics tools and logistics processes, can be fully digitized [7]. This requires the widespread use of the Internet of things, mobile Internet and other advanced technologies to collect relevant logistic data for all scenarios and operations [8]. The digitization of logistics elements is simply the digitization of goods, packaging, logistics documents and personnel. Digitization of logistics facilities is the digitization of fleets, terminals, cargo terminals, warehouses, shelves, sorting and transportation systems, etc. Digitization of logistics tools includes digitization of operational tools such as vehicles, forklifts, pallets, stackers and so on [9]. Digitization of logistic operations refers to the collection of data from the person who passes them, transfer vouchers, loaded goods, vehicle loading time, etc. in other words, collecting information about the entire process, from start to finish download [10], [11].

Based on the application of new information technologies in the field of logistics, new logistics concepts are being formed, such as Party Logistics.

The concept of PL (Party Logistics) is based on determining the level of involvement of independent companies (logistics providers/operators) to solve business problems in the interests and on behalf of the customer (manufacturer, distributor, etc.).

Currently, 1PL-, 2PL-, 3PL-, 4PL- and 5PL-logistics are distinguished.

As a rule, 1PL is autonomous logistics, when all necessary operations (warehousing, transportation, etc.) are carried out by the cargo owner independently using his own infrastructure and personnel.

2PL (Second Party Logistics) is the simplest form of logistics outsourcing: a specialized company, within the framework of contractual obligations, takes part on the tasks of transporting goods and the technical management of inventory.

3PL (Third Party Logistics) is a more developed form of outsourcing: in addition to typical tasks, a professional logistics company can carry out transshipment, sorting, packaging and other processing of goods, as well as provide the client with other additional services with a high share of added value (including using subcontractors). In this case, the full range of logistics services, from delivery and address storage to order management and tracking the movement of goods, is transferred to the transport and logistics organization. However, its main task is not to work with the entire supply chain, but only to perform qualitatively a certain set of physical operations.

Since the majority of 3PL companies focus mainly on fulfilling tasks, rather than on the process as a whole, as the concept of supply chain management requires, this led to the emergence of the next level of outsourcing - 4PL (Fourth Party Logistics).

4PL is an integrator company that accumulates the resources, capabilities and technologies of its own organization and other enterprises (usually 3PL providers) for the design, creation and support of integrated supply chain management solutions. [8] Unlike a 3PL company with a wide range of services, at the 4PL level, the emphasis is on the analysis and re-engineering of business processes in the client company and the introduction of technologies in the interests of the entire supply chain. That is, the proposed 4PL solutions are strategic in nature, and the services of 3PL providers are tactical. An example of a 4PL intermediary is FedEx, which provides its customers from 215 countries with solutions related to transportation and document management within the supply chain.

The latest developments in the field of network connections and intelligent database software, together with the development of outsourcing and strategic partnerships among leading multinational companies, led to the emergence in the mid-2000s. a new level of logistics outsourcing - 5PL (Fifth Party Logistics).

The activities of 5PL providers are supported by modern network computer technologies. Like 4PL, 5PL do not possess tangible assets and are aimed at strategic supply chain management but are more focused on the “virtual enterprise” model.

A virtual enterprise is understood as “a dynamic open business system based on the formation by legally independent enterprises of a single information space for the purpose of sharing their technological resources for the implementation of all stages of the project (customer order) from sources of primary raw materials to delivery of products to the final consumer”.


*Conference’10*, Month 1–2, 2010, City, State, Country.

DOI: http://dx.doi.org/10.1145/12345.67890

# 3. DATA SOURCES

Big data analytics is applicable at all levels of the supply chain management of a logistics company, providing consolidation and interchange of data between levels. Information flows of data in them, according to the authors Normunds Krumins, Kalvis Vitolins [12] is formed from the following sources: 1) customers; 2) data available to the enterprise; 3) public information - social networks, mobile applications and loyalty programs used by consumers; 4) information from consultants, enterprise management, specialist research, etc. In addition, as a result of the analysis and the software product development process, it turned out that a supplier should be added to this list as one of the types of customers, but with its own list of data.

Clients are a source of very valuable information, for example, in the form of orders. This allows you to get information about the most sought-after goods and seasonal demand, and thanks to this, more accurately plan the volume of trade. Similarly, they also receive information on the frequency and volume of orders, transportation costs, etc. Customer surveys are also very important, which allow you to obtain information that is urgently needed.

The data available to the enterprise are, for example, accounting reports, the results of internal research, etc. Usually, these data are not distributed according to the needs of the logistics system, and most often this information is mediated.

Public information mainly means the publication of the results of various surveys, studies, information provided by public organizations (for example, chambers of commerce and industry) and professional associations. In addition, social networks, mobile applications and consumer loyalty programs serve as sources of data on consumers and their preferences. Thanks to these data, it is possible to more accurately forecast demand, develop an improved inventory management program, analyze user activity and interests in order to focus the product or service through contextual advertising.

Another important source is the so-called special information, which is provided by both the management of the enterprise and various consultants and research performers on order.

Suppliers of goods now collect data from various sources, including the origin of the goods, point of arrival, current location. In addition, the delivery itself serves as a data source, providing information about the route, location, information coming from GPS navigators, the provider, and the optimality of the path.

Another class of data sources that is involved in the supply chain is thousands of stores, commodity items, links in the supply chain, millions of customer customers, countless financial and information transactions. And here, big data analytics opens new possibilities.

For example, table 1 presents the variety of data in the supply chain, considering the results of the analysis of our own work and according to the report of the logistics company DHL [13] and [14].

**Table 1. Supply chain data variety**

| Data Source Area | Data | Frequency |
|---|---|---|
| Sales | Price, quality, quantity of goods sold, time, place and date of sale, customer information | Daily hourly |
| Consumers | Information about decisions and buying behavior (which products I searched for, which ones I bought, frequency of purchases, cost and time) | Information comes from every click, bank transfer |
| Stocks | Not aggregated information from multiple locations | Hourly |
| Place and time | Sensory data for determining the place of goods in a store, in a warehouse, in a distribution center, in transport | Constant updates of goods movements |
| Supply | Touch data to determine the best way to deliver goods, weather data, transport information | Hourly |
| Suppliers | Information about the supplier, its location, behavior (which goods it delivers, what distances, the ability to deliver goods across borders, by land, land or air, cost and time) | Daily |
| Marketing | Demand analysis | Historical data, blogs, news, ratings, logs, call center records, consumer surveys, social networks | Demand forecasting and related inventory management program |
| Demand forecasting | Service Level Improvement | Historical data on orders, media data, data from social networks. | Preference analysis |
| Procurement | Negotiations with suppliers | Financial data on the supplier, capacity and customers | contextual advertising |
| Inventory Management | Lower storage costs | Transportation data, warehouse capacity data, loading data | Increased vehicle recycling; |
| Production | Troubleshooting | Internet of things, historical data | Decrease in time of provision of transport; |
| Transportation | Real-time route optimization | Traffic density, weather conditions, transport system restrictions, GPS data | Increase customer satisfaction. |
| Transport management | Reduced repair costs | Technical condition of the car, fuel consumption. | Open negotiations |

In terms of procurement, where big data sources are an incredible amount of information and financial interactions with suppliers, the use of big data analytics allows you to make relationships with suppliers more open and reliable, due to greater disclosure of the general characteristics of companies and financial information. Given the globalization of supply chains, this opportunity is attractive, including for a collaboration of companies.

Finally, thanks to data from GPS sensors and cameras installed on the streets, the cargo transportation function also can the ability to be optimized both in terms of cargo delivery routes and timetables, and in terms of personnel management.

Table 2 shows some examples of problem solving at different levels of supply chain management, which can be solved using big data analytics based on data [15].

**Table 2. Examples of solving logistic problems considering account data sources**

| Function | Problem | Data types | Decision |
|---|---|---|---|

# 4. BIG DATA PROCESSING METHODS

Big data analytics for logistics offers cost reduction methods, an optimal pricing strategy, a strategy to optimize logistics processes and greatly facilitates the decision-making process.

There are three areas of logistics activity, the development of which is facilitated by the Big Data technology [17]:

1) Improving the efficiency of operations, especially the “last mile”, the final stage of the supply chain, which is the costliest. Therefore, logistics systems focus on this issue. The well-known DHL delivery service has divided the solution to this problem in two stages. The first is the optimization of the route of traffic in real time. The second implies a new delivery model in which people who are not working in the company but moving along the route necessary for the company are used. This is done using Big Data technology such as complex event processing and geocorrelation through a special application. This solution is much more efficient than the planned and distributed workforce.

2) Improving the quality of customer service associated with the need to monitor the level of consumer demand and satisfaction. Data obtained and processed using Big Data technology and Web mining can reduce the loss of the customer base and better understand consumer demand.

3) Implementation of an effective and/or new business model related to the analysis of global supply data. Logistics is an indicator of the macroeconomic situation. Transportation of goods

around the world is an important factor in the development of the economy. The types and volumes of goods delivered to various regions make it possible to determine their consumer demand. By receiving and analyzing global supply data, logistics companies can obtain the extensive and detailed information necessary to determine the most effective business model.

Logistics Big Data technology is implemented by supply chain management systems (SCM) in integration with customer relationship management systems (CRM). These systems manage data such as customer transactions, inventory, advertising, customer relationships, their preferences, sales management infrastructure, financial data, etc. The enterprise resource planning system (ERP) is used as a data source.

SCM uses quantitative and qualitative methods to predict the various consequences and solve supply chain problems. The use of GPS and radio frequency identification data (RFID), sensors on storage equipment help to accurately track the status of the goods. SCM with Big Data technology is usually supermarket chains. SCM solves many problems of the enterprise, especially inventory, allowing you to see the exact number of goods on store shelves at a certain point in time.

Consider some of the Big Data features used in logistics:

1) Supply Volume Analysis is the ability to forecast supply volumes on a specific day of the week, month, year is necessary for optimal budget allocation. Analysis of existing data will help to predict peaks in production volumes and generate recommendations for optimizing these processes.

2) Data on goods with special storage and transportation conditions, i.e. some products require special storage conditions. For example, some products and medicines must be stored at a certain temperature, fragile items require special transportation conditions. The transportation of such goods in the end turns out to be expensive both for the logistics company and for the client, especially in the case when, due to incorrect storage and transportation conditions, the goods are damaged. Here the technology of the Internet of things come to help, for example, sensors scan temperature conditions, stock levels, etc. Data is analyzed offline by generating information about the safest and most economical ways of transporting and placing goods.

3) Economy route where Big Data analysis will reveal the most reliable and economical routes. The output will help you choose the best airline or warehouse company.

4) Risk analysis or risk assessment will significantly help reduce company costs. Big Data predicts various unforeseen situations and offer alternative strategies to help them avoid or ways to minimize losses.

5) Delivery time analysis is always important for customers to know the exact delivery date, which is very difficult to predict. Delivery time depends on various factors: the number of orders, goods, the situation on the road, the state of transport, etc. By analyzing the full volume of this data, an approximate delivery time can be assumed.

6) Real-time data analysis will provide accurate information about the current location of the product.

7) Web analysis, i.e. Big Data analysis will help in providing a more personalized catalog of services to website visitors. Depending on the interests of the visitor, different services can be automatically offered to him.

## 5. ANALYTICAL MODEL OF THE LOGISTIC PROBLEM

The development of analysis and visualization technologies allows companies to effectively use analytical models. Analytical models allow companies to quickly make effective decisions based on effective analysis of huge data sets [16].

Instead of the traditional process of chaotic collection of operational data and spontaneous reporting, it should be replaced by an advanced approach to collecting data from various data sources using analytical models aimed at specific functional areas.

Studies, according to Deloitte [16], DHL [13], show that companies are beginning to understand the criticality of big data analysis for the supply chain

Key factors for the development of analytical models:

- Data dissemination;
- Available data warehouses;
- Data processing speed;
- Internet development;
- Hi-tech;
- Visualization features.

Advantages of using analytical models:

- The optimal distribution of stocks in warehouses, considering the associated costs and demand.
- Simulation of demand depending on key factors.
- Consumption forecast based on analysis of customer behavior.

The most complete study of the process of functioning of the system can be carried out if explicit dependencies are known that connect the desired characteristics with the initial conditions, parameters, and variable systems. However, such dependencies can only be obtained for relatively simple systems. When systems become more complicated, their analysis using analytical methods encounters certain difficulties, which is a significant drawback of the method. In this case, to use the analytical method, it is necessary to simplify the initial model to be able to study the general properties of the system.

The main stages of the development of the analytical model are presented in Figure .1. It should be noted that the decision-making model can be determined depending on the functional area of application.

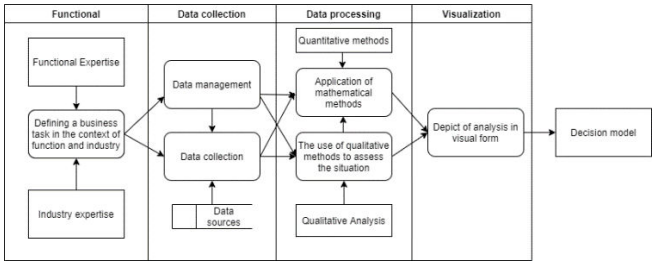


**Figure 1 - Analytical model adopted from [16]**

## 6. CONCLUSION

Thus, we can conclude that Big Data technology with combination of the Internet of things, can support logistics in improving the efficiency of operations and the quality of customer service and receiving and analyzing global supply data.

Generation of large amount data implies the development of analytical models. Analytical models allow logistics companies to quickly make effective decisions based on effective analysis of huge data sets. In this paper we propose an analytical model which depicts main stages of Big Data analysis: functional expertise, data collection, data processing and visualization. In addition, Big Data allows automated logistics systems to function through intelligent routing of many different data sets and data streams.

## 7. ACKNOWLEDGMENTS

This work was carried out as a part of the research project of grant financing of the Ministry of Education and Science of the Republic of Kazakhstan BR05236517-OT-18 “Platform for the digital transformation of the business processes of the national economy”.

## 8. REFERENCES

[1] Fernie J., Sparks L. 2014. Logistics and retail management: emerging issues and new challenges in the retail supply chain.

[2] Barreto L, Amaral A, Pereira T. 2017. Industry 4.0 implications in logistics: an overview. Procedia Manuf 13:1245–1252.

[3] Kawa A. 2017. Fulfillment service in e-commerce logistics. LogForum 13:429–438.

[4] Aliresearch. 2017. Big data development report on China smart logistics. http://i.aliresearch.com/img/20170328/20170328173815.pdf

[5] Manyika J., Chui M., Brown B., Bughin J., Dobbs R., Roxburgh C., Byers A.H. 2011. Big data: the next frontier for innovation, competition, and productivity. McKinsey Global Institute.

[6] Walker J.S. 2014. Big data: a revolution that will transform how we live, work, and think. Int J Advert 33:181–183.

[7] Yang M., Mahmood M., Zhou X., Shafaq S., Zahid L. 2017. Design and implementation of cloud platform for intelligent logistics in the trend of intellectualization. China Comm 14:180–191.

[8] Kataria M., Mittal M.P. 2014. Big data: a review. Int J Comput Sci Mobile Comput 3:106–110.

[9] Zhong R.Y., Newman S.T., Huang G.Q., Lan S. 2016. Big data for supply chain management in the service and manufacturing sectors: challenges, opportunities, and future perspectives. Comput Ind Eng 101:572–591.

[10] Sagiroglu S., Sinanc D. 2013. Big data: a review. 2013 International Conference on Collaboration Technologies and Systems (CTS) IEEE:42–47.

[11] Ghosh D. 2015. Big data in logistics and supply chain management: a rethinking step. 2015 International Symposium on Advanced Computing and Communication (ISACC) IEEE:168–173.

[12] Krumin'sh N., Vitolin'sh, K., Informatsionnyye sistemy v logistike, [Online]. Available: http://transportinform .com/ logistika/585-informaczionnye-sistemy-v-logistike.html. [Accessed 8 апрель 2020].

[13] Kuckelhaus, M., Zeiler, K., Jeske, M., Gruner, M., Weis, F., "BIG DATA IN LOGISTICS. A DHL perspective on how to move beyond the hype," DHL Customer Solutions & Innovation Represented by Martin Wegner, December 2013. https://www.detecon.com/sites/default/files/DHLDetecon_Big_Data_Whitepaper.pdf.

[14] Poleshuk Ye. N. 2017. Mekhanizm formirovaniya ratsional'noy sistemy logisticheskogo obespecheniya raspredelitel'nykh setey, in Perspektivy razvitiya logistiki i upravleniya tsepyami postavok, Moskva, 536-550

[15] Rozados I., Tjahjono B. 2014. Big Data Analytics in Supply Chain Management: Trends and Related Research," in *6th International Conference on Operations and Supply Chain Management*, DOI:10.13140/RG.2.1.4935.2563

[16] Gasparyan V., Kalinkin K. 2014. "Big Data" and Analytical Models in Logistics and SCM", Deloitte.

[17] DHL: website. 2017. URL: http://www.dhl.com/content/dam/downloads/g0/about_us/innovation/CSI_Studie_BIG_DATA.pdf.

[18] Shadi Aljawarneh, Muneer Bani Yassein, and We’am Adel Talafha. 2018. A multithreaded programming approach for multimedia big data: encryption system. Multimedia Tools Appl. 77, 9 (May 2018), 10997–11016. DOI:https://doi.org/10.1007/s11042-017-4873-9

[19] Vangipuram Radhakrishna, Shadi A. Aljawarneh, Puligadda Veereswara Kumar, and Kim-Kwang Raymond Choo. 2018. A novel fuzzy gaussian-based dissimilarity measure for discovering similarity temporal association patterns. Soft Comput. 22, 6 (March 2018), 1903–1919. DOI:https://doi.org/10.1007/s00500-016-2445-y

[20] Shadi Aljawarneh, Shadi Aljawarneh, and Manisha Malhotra. 2017. Critical Research on Scalability and Security Issues in Virtual Cloud Environments (1st. ed.). IGI Global, USA.

[21] Shadi Aljawarneh, Muneer Bani Yassein, and We’am Adel Talafha. 2017. A resource-efficient encryption algorithm for multimedia big data. Multimedia Tools Appl. 76, 21 (November 2017), 22703–22724. DOI:https://doi.org/10.1007/s11042-016-4333-y

[22] Muneer Bani Yassein, Shadi Aljawarneh, and Esraa Masadeh. 2017. A new elastic trickle timer algorithm for Internet of Things. J. Netw. Comput. Appl. 89, C (July 2017), 38–47. DOI:https://doi.org/10.1016/j.jnca.2017.01.024

**Columns on Last Page Should Be Made As Close As Possible to Equal Length**

# Authors' background

| Your Name | Title* | Research Field | Personal website |
|---|---|---|---|
| Aiman Moldagulova | Associate professor | Information Systems | https://www.iitu.kz/ru/departaments/profshow/12 |
| Ryskhan Satybaldiyeva | Associate professor | Information Systems | https://www.iitu.kz/ru/departaments/profshow/11 |
| Abu Kuandykov | Full Professor | Information Systems | https://www.iitu.kz/ru/eduprog/show/45 |

***This form helps us to understand your paper better, the form itself will not be published.**

***Title can be chosen from: master student, Phd candidate, assistant professor, lecture, senior lecture, associate professor, full professor**