

Использование Big Data в России

Попова Дарья Григорьевна, студентка 4-го курса финансового факультета Российского экономического университета имени Г.В. Плеханова, г. Москва, Российская Федерация

Email: dassha.96@mail.ru

Аннотация

Это статья раскрывает сущность больших данных и их роль в современном бизнесе. На примере ключевых компаний финансового и транспортного сектора, а также сектора розничной продажи, мы можем проследить развитие концепта Big Data в России, и найти ответ на вопрос «Есть ли будущее у Big Data в России?»

Ключевые слова: финансы, Big Data, аналитика.

Big Data in Russia

Popova Darya Grigorievna, student, Russian Plekhanov University of Economics, Moscow, Russian Federation

Email: dassha.96@mail.ru

Annotation

This article reveals the essence of Big Data, their role in modern finance. Based on the example of key companies in the financial, transportation and retail sector it is possible to trace the development of the Big Data concept in Russia and find the answer of the question “Does Big Data have future in Russia?”

Keywords: finance, Big Data, analytics.

In recent years, much more data is available to the world than we could have imagined decades ago. Google stopped reporting how much data they store in 2010. At that time, it was 100 PBs. YouTube orders are measured in Exabyte, and the number of downloadable video hours per minute exceeded 72 hours. The number of Facebook users has exceeded a billion people in August 2012. Twitter posts about 124 billion tweets per year, an average of 4,500 per second. Companies produce terabytes of information.

The main tasks for these companies today are to process, analyze and use the processed data to obtain wise decisions and improve the efficiency of the company as a whole.

Big Data is the new generation of technologies designed to cost-effectively extract useful information from very large volumes of diverse data through the high-speed collection, processing and analysis. At the same time, storage technologies (relational DBMS with support for massively parallel processing, NoSQL, Hadoop) are less important than workload parameters such as data volume, transfer speed, heterogeneity and value.

The main question of the application of Big Data technologies in the financial sector is will it increase productivity, avoid risks, predict and expand the capabilities of companies.

Big Data is relatively new for Russian business, but it has become widespread in Western countries. The main difference between Big Data and traditional database is represented in the Table 1.

Table 1. Difference between Big Data and traditional database.

Characteristics	Big Data	Traditional Database
Volume of information	From gigabytes (10^9 bytes) to terabytes (10^{12} bytes)	From petabytes (10^{15} bytes) to exabytes (10^{18} bytes)
Storage method	Centralized	Decentralized
Structuredness	Structured	Semi-structured
Data Storage and Processing Model	Vertical model	Horizontal model
Relationship between data	Strong	Weak

It is impossible to analyze Big Data without software and equipment.

SQL, which is a structured query language that allows to create and modify data. The corresponding database management system (DBMS) is responsible for managing the data array.

NoSQL stands for Not Only SQL (not just SQL), includes a number of approaches aimed at implementing a database that differs from the models used in traditional, relational DBMSs. They are convenient for use with an ever-changing data structure. For example, to collect and store information in social networks.

MapReduce is model of the distribution of calculations, Hadoop is created to implement the search and context mechanisms of highly loaded sites (Facebook, eBay) and SAP HANA is a high-performance NewSQL platform for data storage and processing.

Analysis of Big Data is associated with all types of human activities. Nevertheless, in this article, the special attention is paid on the use of Big Data in Business and Finance.

The modern financier is more and more associated with a person who can handle a huge amount of information and has strong analytical skills. The change in paradigms has meant that today's managers and economists should work not only with large volumes of information but also with new types of data, in order to achieve more accurate and efficient financial solutions.

The field of use of Big Data technologies in business is extensive [3]. So, with the help of Big Data, you can learn about the preferences of customers, the effectiveness of marketing campaigns or conduct a risk analysis. The results of the survey of the IBM Institute, on the directions of using Big Data in companies is the following: 53% of companies use the Big Data in the customer service area, the second most popular direction is operational efficiency (40%), in the field of risk management, Big Data is less common at the moment (7%).

Big Data has become widespread in many business sectors [1]. They are used in healthcare, telecommunications, trade, logistics, financial companies, and public administration.

For instance, in the databases of retail stores, a lot of information about customers, inventory management system, supply of marketable products can be accumulated. This information can be useful in all areas of the shops.

So, with the help of the accumulated information, it is possible to manage the supply of goods, its storage and sale [11]. Based on the collected information, it is possible to predict the demand and supply of goods. Also, the data processing and analysis system can solve other problems of the retailer, for example, optimize costs or prepare reports.

One of the first offline stores that implemented the analysis of the Big Data in Russia was Lenta. With the help of Big Data, retail began to study information about customers from cash vouchers. The retailer collects information for compiling behavioral models, which makes it possible to make decisions more rationally at the level of operational and commercial activities.

In the telecommunications industry, large numbers of data are received from mobile operators. Cellular operators have one of the most extensive databases, which allows them to conduct the most profound analysis of the accumulated information.

The main purpose of data analysis is to retain existing customers and attract new ones [9]. To do this, companies conduct segmentation of customers, analyze their traffic, determine the social affiliation of the subscriber. In addition, technologies are used to prevent fraudulent financial transactions.

In the field of transportation services and logistics, Big Data technologies also play an important role [4]. With the help of data analysis, you can analyze the route taking into account the time and fuel

costs, calculate the optimal path, and most quickly and efficiently process the incoming application from the client.

RZD has implemented Big Data technologies in conjunction with SAP. These technologies helped reduce the period of reporting by 43.5 times (from 14.5 hours to 20 minutes), improve the accuracy of cost allocation by 40 times. Big Data was also introduced into planning and tariff regulation processes.

Big data are used both for the extraction of useful minerals and for their processing and marketing. Based on the information received, enterprises can draw conclusions about the efficiency of field development, track the schedule of capital repairs and equipment condition, predict the demand for products and prices.

A significant proportion of Big Data users are professionals from the financial industry [2]. One of the successful experiments was carried out in the Ural Bank for Reconstruction and Development, where the information base was used to analyze clients, the bank began offering specialized loan offers, deposits and other services. During the year of using these technologies, the company's retail loan portfolio grew by 55%.

Alfa-Bank analyzes information from social networks, process applications for a loan, analyze the behavior of visitors to the company's website.

Sberbank also began processing the array of data to segment customers, prevent fraud, cross-selling and risk management. In the future, it is planned to improve the service and analyze the actions of clients in real time.

The Russian Regional Development Bank analyzes the behavior of plastic card holders. This makes it possible to identify atypical operations for a particular client, thereby increasing the likelihood of detecting money theft from plastic cards [8].

Because of the introduction of a special system for analyzing the large data of its borrowers, VTB 24 last year increased the net interest margin from credit operations by 16%, and this brought about 19 billion rubles net interest income [10].

Previously, "VTB 24" divided customers into nine marketing groups, and now introduced for each seven risk profiles. It turned out 63 segments. Having improved segmentation, the bank began to offer customers individual conditions, depending on the solvency. Less risky borrowers are offered products with lower rates - and vice versa. [6] The cost of the system for three years amounted to about \$ 33 million and it has already paid off 10 times.

Algorithmic trading has become synonymous with big data due to the growing capabilities of computers. The automated process enables computer programs to execute financial trades at speeds and frequencies that a human trader cannot. [5] Within the mathematical models, algorithmic trading provides trades executed at the best possible prices and timely trade placement, and reduces manual errors due to behavioral factors.

According to the results of the research of CNews Analytics and Oracle, the level of maturity of the Russian market Big Data has increased over the past year. The main drivers for this process has become:

Increased interest on the part of users to the possibilities of Big Data as a way to increase the competitiveness of the company [7]:

- Development of methods for processing media files on the global level;
 - Transfer of servers processing personal information to the territory of Russia, in accordance with the adopted law on the storage and processing of personal data;
 - Implementation of the sectoral plan for import substitution of software. This plan includes state support for domestic software producers, as well as the provision of preferences for domestic IT products in procurement at public expense;
 - In the new economic situation, when the dollar rate almost doubled, there will be a trend towards increasing use of services of Russian providers of cloud services, rather than foreign ones;
 - Creation of technoparks that contribute to the development of the information technology market, including the Big Data market;
 - State program for the introduction of grid-systems, based on the technology of Big Data.
- The main barriers to the development of Big Data on the Russian market are:
- Ensuring the security and confidentiality of data;
 - Lack of qualified staff;
 - Lack of accumulated information resources to the level of Big Data in most Russian companies;
 - Difficulties in introducing new technologies into established information systems of companies;
 - High cost of Big Data technologies, which leads to a limited number of enterprises that are able to implement these technologies;
 - Political and economic uncertainty that led to the outflow of capital and the freezing of investment projects on the territory of Russia;

- The growth of prices for imported products and a surge in inflation, according to IDC, hamper the development of the entire IT market.

As of today, the Russian market of Big Data is not so popular as in developed countries. However, examples of Lenta, Yandex, Sberbank and other companies have revived the market's interest for new technologies. Analysts have quite optimistic forecasts regarding the Russian market. IDC believes that the share of the Russian market over the next 5 years will increase, in contrast to the market of Germany and Japan.

In 2017, the world's revenue on the Big data market reached \$150.8 billion, which is 12.4% more than last year.

By 2020, the volume of Big Data of Russia will grow from the current 1.8% to 2.2% of the global data volume. The amount of information will grow, according to EMC data up to 980 exabytes in 2020.

Список использованных источников

1. Бабурин В.А., Яненко М.Е. Технологии Big Data в сервисе: новые рынки, возможности и проблемы. – 2014. – № 1 (27). – С. 100-105.
2. Блинов А.О. Управление организациями с помощью когнитивных технологий / Блинов А.О., Рахманкулов И.Ш. // Экономика и управление в отраслях и сферах деятельности. – 2015. – № 4. – С.33-39.
3. Большие данные принесли «ВТБ 24» в прошлом году 19 млрд рублей дополнительно // Газета «Ведомости» [электронный ресурс] – Режим доступа. – URL: <https://www.vedomosti.ru/technology/articles/2017/05/05/688849-bolshie-dannie-vtb-24> (дата обращения 15.04.2018).
4. Быкова А.А., Морковкина Е.В. Капитал отношений как фактор повышения стоимости компании // Корпоративные финансы. – 2013. – №4 (28). – С. 19-36.
5. Волкова Ю.С. Большие Данные в современном мире // Концепт. –2016. – Т. 11. – С. 1171–1175.
6. Когнитивное унифицированное управление конечными точками: будущее мобильного управления и безопасности // Официальный сайт IBM [электронный ресурс] – Режим доступа. – URL: <https://www.ibm.com/security/mobile/maas360> (дата обращения 15.04.2018).
7. Лофран Т., Макдональд Б. Измерение удобочитаемости в раскрытии финансовой информации. – 2014. – № 4. – С. 1643-1671.

8. Макеева Е.Ю., Аршавский И.В. Применение нейронных сетей и семантического анализа для прогнозирования банкротства. Корпоративные финансы. – 2014. – № 8 (4). – С. 130-141.
9. Мальцева С.В., Лазарев В.В. Маркетинговая аналитика в сфере электронного бизнеса на основе больших данных // Информационные технологии в проектировании и производстве. – 2015. – № 1. – С. 62-67.
10. Санджив Д. Матричные показатели: системная оценка системного риска. – 2016. – № 4. – С. 33-51.
11. Polonski V. The evolution of social networking sites: the rise of content-centric platforms which favour the perpetual present [электронный ресурс] – Режим доступа. – URL: <http://blogs.lse.ac.uk/impactofsocialsciences/2014/01/15/the-evolution-of-social-network-sites-in-2014> (дата обращения 14.04.2018).

References

1. Baburin V.A., Yanenko M.E. Tekhnologii Big Data v servise: novye rynki, vozmozhnosti i problem, 2014, No. 1 (27), pp. 100-105.
2. Blinov A.O. Upravlenie organizatsiyami s pomoshch'yu kognitivnykh tekhnologii / Blinov A.O., Rakhmankulov I.Sh. // Ekonomika i upravlenie v otraslyakh i sferakh deyatelnosti, 2015, No. 4, pp. 33-39.
3. Bol'shie dannye prinesli «VTB 24» v proshlom godu 19 mlrd rublei dopolnitel'no // Gazeta «Vedomosti»
<https://www.vedomosti.ru/technology/articles/2017/05/05/688849-bolshie-dannye-vtb-24>
4. Bykova A.A., Morkovkina E.V. Kapital otnoshenii kak faktor povysheniya stoimosti kompanii // Korporativnye finansy, 2013, No. 4 (28), pp. 19-36.
5. Volkova Yu.S. Bol'shie Dannye v sovremennom mire // Kontsept, 2016, Vol. 11, pp. 1171-1175.
6. Kognitivnoe unifitsirovannoe upravlenie konechnymi tochkami: budushchee mobil'nogo upravleniya i bezopasnosti // Ofitsial'nyi sait IBM
<https://www.ibm.com/security/mobile/maas360>
7. Lofran T., Makdonal'd B. Izmerenie udobochitaemosti v raskrytii finansovoi informatsii, 2014, No. 4, pp. 1643-1671.

8. Makeeva E.Yu., Arshavskii I.V. Primenenie neironnykh setei i semanticheskogo analiza dlya prognozirovaniya bankrotstva. Korporativnye finansy, 2014, No. 8 (4), pp. 130-141.
9. Mal'tseva S.V., Lazarev V.V. Marketingovaya analitika v sfere elektronnoho biznesa na osnove bol'shikh dannykh // Informatsionnye tekhnologii v proektirovanii i proizvodstve, 2015, No. 1, pp. 62-67.
10. Sandzhiv D. Matrichnye pokazateli: sistemnaya otsenka sistemnogo riska, 2016, No. 4, pp. 33-51.
11. Polonski V. The evolution of social networking sites: the rise of content-centric platforms which favour the perpetual present
<http://blogs.lse.ac.uk/impactofsocialsciences/2014/01/15/the-evolution-of-social-network-sites-in-2014>