УДК 338.1

BIG DATA IN MODERN BUSINESS

N.A.Naim
assistant at the Tashkent University of Information Technologies named after Muhammad al Khorazmiy, e-mail:nodira343y@mail.ru

N.A.Naim
She graduated from the Tashkent University of Information Technologies named after Muhammad al Khorazmiy. TUIT teacher on the estate of Muhammad al Khorazmiy

Abstract. This article discusses the main ways to use big data technologies in modern business. Also, in what cases can Big data technologies be used in the modern market and why big data is gaining more and more popularity despite the high cost.

Keywords: big data, Hadoop, DB, volume, velocity, variety, marketing.

It's no secret that today the amount of data that needs to be stored and processed is growing exponentially. For example, the amount of data stored on the Internet is increasing by about 40% annually. Interestingly, on the one hand, it is the development of modern information technologies that allows and contributes to the fact that the volume of stored and processed data is constantly growing. On the other hand, working with rapidly growing volumes of a wide variety of data requires more and more resources and more complex software solutions. One of the most modern and rapidly gaining popularity technologies is big data. This term began to be used and quickly gain popularity only 8-10 years ago. And today, the world's largest companies, which occupy leading positions in various areas of business activity, are investing billions of dollars in the development of this area. This article discusses the basic concepts associated with big data technology, why some large companies are ready to invest huge funds in the development of this area, while others are not. What are the most common big data-based solutions on the market today, what are the difficulties and what awaits the rest of the IT market with the advent of new technology.

Literally big data means big data. A more detailed definition can be formulated as follows. Big data is a series of approaches and methods for processing a large volume and a significant variety of data that are difficult to process in conventional ways. The purpose of big data processing is to obtain new information. In this case, the data can be both processed (structured) and fragmented (that is, unstructured). If we talk about the types of data for which the processing methods in question are applicable, then this can be completely different information: documents, blogs, social networks, any client data, or even information about the actions performed by clients. Also information coming from measuring devices, etc. But this is all mostly textual information. In addition, audio and video data, images, etc. can be processed.

Variety (from English variety). The data may not be uniform. That is, they can be fragmented, structured, unstructured or partially structured. And the point is to process different types of data at the same time. Also, to the already considered three V, a fourth is added in different sources. Credibility or credibility (from English veracity). And even the fifth viability or value

(from the English viability or value). In various versions, you can talk about 7V, but three is enough for a basic understanding.

Principles of big data operation Based on the definition of big bata, we can formulate three main principles for working with such data: Horizontal scalability. Since the amount of data is constantly and rapidly growing and information can be arbitrarily large, the system that involves the processing of this data must be extensible. For example, if the amount of data has grown 2 times, then it should be possible to increase the hardware capacity by 2 times in the cluster and the system will continue to work without performance losses. Fault tolerance. The principle of horizontal scalability discussed above implies that there can be many machines in a cluster. For example, at Yahoo, the cluster has over 40,000 machines. At the same time, it is assumed that some of these machines will regularly fail. Big data practices must take into account the likelihood of such failures and keep the system running without any significant consequences.

Data locality. In large distributed systems, the data in use is stored on a large number of machines. But if the data is physically located on one server and processed on another, then the resources required to transfer the data may exceed the costs of processing the data. Therefore, when designing solutions based on big data, one of the most important principles is the principle of data locality, the essence of which is that the data is processed and stored on the same machine.

Features of application and role in modern business Studying the variety of modern technologies for storing and processing data, a logical question arises. What are the methods and approaches called big data invented for? What is unique about this, how can information processed using these technologies be used, and why are companies ready to invest huge amounts of money in the development of big data? First, unlike big data, conventional databases (DBs) cannot store and process such huge amounts of data (hundreds and thousands of terabytes). And it's not even about analytics, but only about data storage. In the classical sense, the database is designed for fast processing (storage, modification) of relatively small amounts of data or for working with a large stream of small records, i.e., a transactional system. With the help of big data, this main task is solved - the successful storage and processing of large amounts of data.

Secondly, in big data, heterogeneous information is structured that comes from various sources (images, photos, videos, audio and text documents) into one single, understandable and acceptable form for further work. Thirdly, in big data, analytics are formed and accurate forecasts are made based on the received and processed information. What is it for and where can it be applied in practice? For clarity and in order to formulate the answer in simple words, let's look at the example of typical business tasks in marketing. Possessing such information as:

– a complete understanding of your company and business, including in terms of statistical information and figures;

- detailed information about competitors;

– new and detailed information about their customers;

- all this will allow you to succeed in attracting new customers, significantly increase the level of service provided to current customers, better understand the market and your competitors, and therefore get ahead by dominating them. Considering the above results that big data can achieve, it explains the desire of companies trying to conquer the market to invest in these modern data processing methods today in order to get increased sales and reduced costs tomorrow. And more specifically, then:

– increase in additional sales and cross-sales due to a better knowledge of customer preferences; - search for popular products and reasons

- why they are bought or vice versa;

- improvement of the provided service or product;

– improving the quality of customer service;

– increasing loyalty and customer orientation;

– fraud prevention (more relevant for the banking sector);

- Reduction of unnecessary expenses. One of the most illustrative and popular examples today, which can be read about in many sources on the Internet, is associated with Apple, which collects data about its users using manufactured devices: phone, tablet, watch, computer. It is because of the presence of such a system that the corporation owns a huge amount of information about its users and subsequently uses it for profit. And there are many such examples today. A brief overview of big data tools Considering the huge volumes of information that need to be stored and processed in the process of work, it should be noted that such manipulations cannot be performed on simple hard drives. And the software that structures and analyzes the accumulated data is a separate intellectual property and in each case is an author's development. At the same time, we can note the most popular tools today, on the basis of which such solutions are created:

- Hadoop & MapReduce;

– NoSQL databases;

– Data Discovery class tools.

An analysis of the features and differences of the listed tools, as well as an overview of solutions that can be offered based on these tools, can be the subject of a separate article. But, as an example, I would like to give a model that, perhaps, is one of the leaders on the market today - this is the Oracle Big Data Appliance X5-2. The estimated cost of such a system in the maximum configuration can reach 30 million rubles per 1 rack. Of course, we are talking about premium industrial systems. Nevertheless, the above example allows us to estimate the order of expenses that will be required for the implementation of such solutions in the company. And this is without taking into account highly specialized specialists and additional IT infrastructure. Therefore, it is not necessary to talk about the use of big data, for example, in small businesses.

Secondly, in big data, heterogeneous information is structured that comes from various sources (images, photos, videos, audio and text documents) into one single, understandable and acceptable form for further work. Thirdly, in big data, analytics are formed and accurate forecasts are made based on the received and processed information. What is it for and where can it be applied in practice? For clarity and in order to formulate the answer in simple words, let's look at the example of typical business tasks in marketing. Possessing such information as:

– a complete understanding of your company and business, including in terms of statistical information and figures;

– detailed information about competitors;

– new and detailed information about their customers;

– all this will allow you to succeed in attracting new customers, significantly increase the level of service provided to current customers, better understand the market and your competitors, and therefore get ahead by dominating them.

Considering the above results that big data can achieve, it explains the desire of companies trying to conquer the market to invest in these modern data processing methods today in order to get increased sales and reduced costs tomorrow. And more specifically, then:

– increase in additional sales and cross-sales due to a better knowledge of customer preferences;

– search for popular products and reasons - why they are bought or vice versa;

– improvement of the provided service or product;

– improving the quality of customer service;

– increasing loyalty and customer orientation;

– fraud prevention (more relevant for the banking sector);

– Reduction of unnecessary expenses. .

Nevertheless, the above example allows us to estimate the order of expenses that will be required for the implementation of such solutions in the company. And this is without taking into account highly specialized specialists and additional IT infrastructure. Therefore, it is not necessary to talk about the use of big data, for example, in small businesses.

References

[1] Alimova F.M., Xaydarova M., Usmonov J., Ta'limda big data va machine learning qo'llanilish tahlili, "Zamonaviy tadqiqotlar, innovatsiyalarning dolzarb muammolari va rivojlanish tendensiyalari: yechimlar va istiqbollar" mavzusidagi Respublika miqyosidagi ilmiy-amaliy konferensiya, Тошкент – 2021. – Б. 241-243.

[2] Алимова Ф.М., Кушманова М.А., Наим Н.А., Перспективы использования больших данных в системе образования, Сборник научных статей VII международной научно-практической конференцияии. "Big data и анализ высокого уровня". Минск– 2021. – С. 193-196.

[3] Алимова Ф.М., Кушманова М.А., Наим Н.А., Big data в медицине электрон, Сборник научных статей VII международной научно-практической конференции. "Big data и анализ высокого уровня". Минск– 2021. – С.221-223.

БОЛЬШИЕ ДАННЫЕ В СОВРЕМЕННОМ БИЗНЕСЕ

Н.А. НАИМ
Преподаватель ТУИТ имении
Мухаммада ал Хоразмий.

Аннотация. В данной статье рассматриваются основные способы использования технологий больших данных в современном бизнесе. Также в каких случаях могут применяться технологии Big data современном рынке и почему большие данные все больше набирает популярность несмотря на высокую стоимость.

Ключевые слова. Big data, Hadoop, БД, объем, скорость, многообразие, маркетинг.