On the application of Big Data technologies in the banking sector

Assoc. Prof. Aliyeva Tarana Abulfaz

Azerbaijan State University of Economics (UNEC),
Department of Information Economy and Technologies
mob.: +99450 3010112 and alilitarana@gmail.com

ABSTRACT

The proposed article is devoted to the application of Big Data technologies in the banking sector. The main characteristics and spheres of application of the Big Data are investigated. The principles of working with large amounts of data are described when using the Big Data technology to collect overdue debts by searching and collecting contact information about the customer, for further establishing contact with him. The article contains a description of the economic effect of the application and significance of Big Data for the implementation of one of the most important functions of the bank. An approach is proposed for creating a multifactor model for client analysis at the stage of assessing the application for a loan to minimize overdue debt.

Key words: Big Data, the traditional database, banking, overdue debt, economic effect

INTRODUCTION

The development of information storage and processing technologies was provoked by the growth of the amount of heterogeneous information. There are such volumes of information that can't be processed by traditional methods, including structured data, media data and random objects. And if with the analysis of the first existing technologies today are less able to cope, then the analysis of the second and third practically remains an excessive work. Studies show that the volume of media data, such as the results of video surveillance, aerial photography, digital medical information, and random objects stored in numerous archives and clouds, is increasing year by year. Particular interest is observed in the direction of using the paradigm of "Big Data". The Big Data concept is a model for large-scale management of arrays of distributed, weakly structured data [1].

The work of foreign scientists, such as James Manyika, Michael Chui, Toporkov VV, Budzko VI, is devoted to the study of Big Data. A significant contribution to the study of this

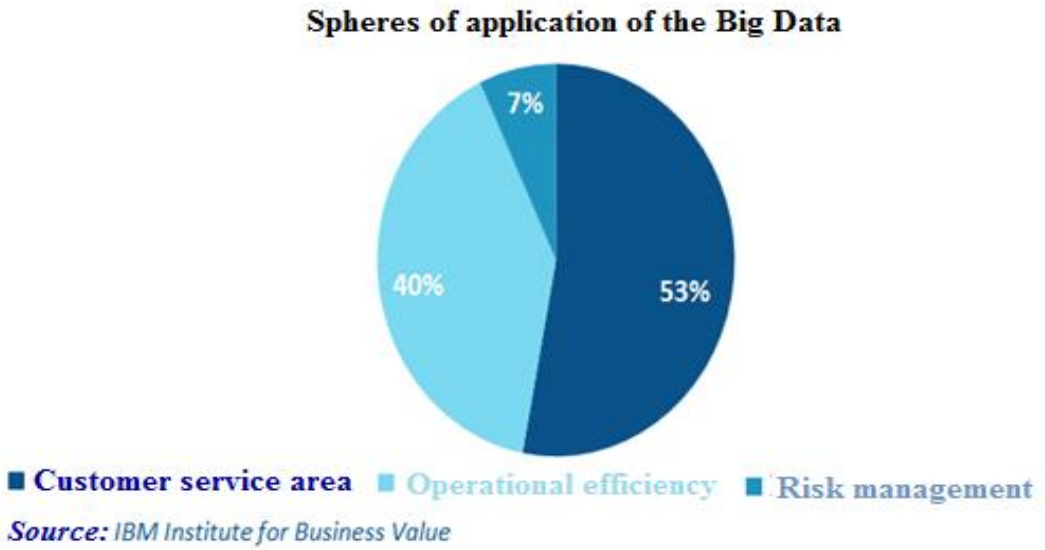
technology is made by major global companies such as: McKinsey & Company, SNews Analytics, SAP, Oracle, IBM, Microsoft, Teradata and many others. They are engaged in data processing and analysis and on the basis of Large data create software and hardware systems. According to the McKinsey Institute report: "Big Data is a set of data, that is beyond the capabilities of typical databases of software tools for data capture, storage, management and analysis" [1]. In essence, the concept of large data means working with information of a huge volume and diverse composition, constantly updated and located in different sources in order to increase work efficiency, create new products and improve competitiveness. Consulting company Forrester gives a brief and fairly clear statement: "Big Data combines techniques and technologies that make sense from data at the extreme limit of practicality" [2].

MATERIALS AND METHODS

Big Data are methods of working with a large amount of structured and unstructured data [3]. Big Data are characterized by three characteristics [3]: 1. Volume - a large amount of data that is constantly increasing 2. Velocity - the speed of working with similar data 3. Variety - the diversity of stored and processed data. Experts from among marketers like to add their "V" here. Someone says more about the veracity, others add that the technology of large data must certainly benefit the business. It is expected that by 2020 the accumulated amount of information on the planet will double every two years. The abundance of data causes a desire to use them for analysis and forecasting. Enormous volumes require appropriate technology.

The field of use of Big Data technologies is extensive. They are now widely used in many business sectors. They are used in healthcare, telecommunications, trade, logistics, financial companies and banks, as well as in public administration.

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. Below are the results of the survey of the IBM Institute, on the directions of using Big Data in companies.



As can be seen from the diagram, most companies use the Big Data in the customer service area, the second most popular direction is operational efficiency, in the field of risk management, Big Data is less prevalent at the moment.

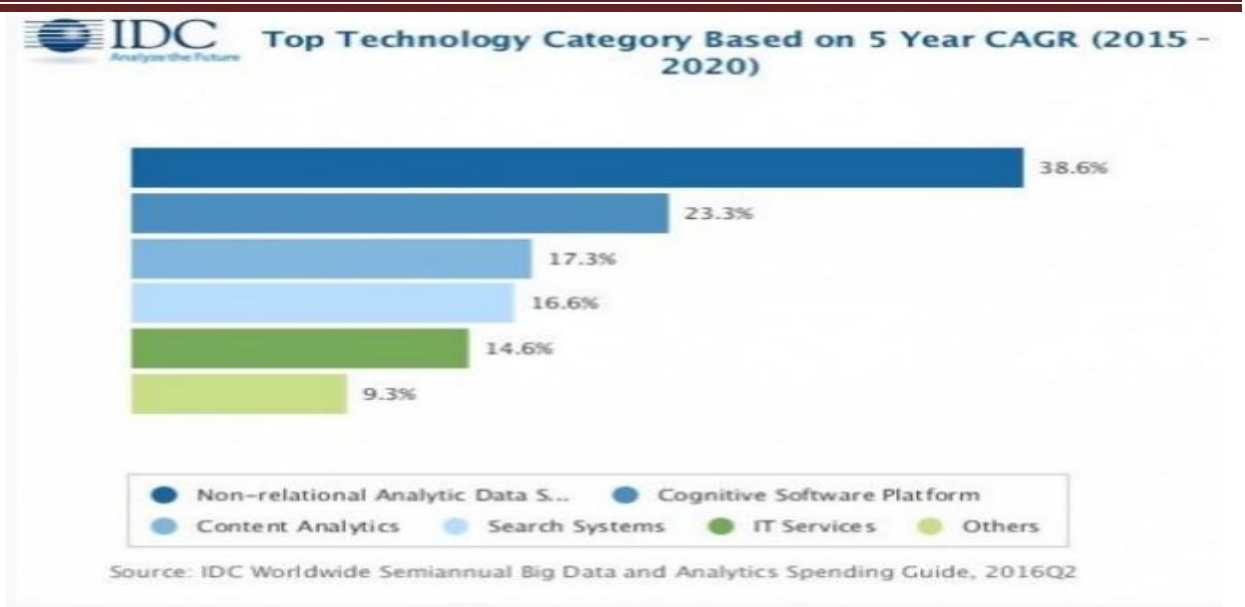
Today, companies need to process a huge amount of data in volumes that are difficult to imagine, which leads to the fact that traditional databases can't cope with such a task, and this leads to the need to implement Big Data technologies. The table presents a comparative characteristic of Big Data and traditional databases. The basis for the formation of this table was the research of the Moscow Stock Exchange [4].

Table 1. Comparative characteristics of Big Data and traditional data

Characteristic	Traditional database	Database of Big Data
Amount of information	From gigabytes (10 ^ 9 bytes) to terabytes (10 ^ 12 bytes)	From petabytes (10 ^ 15 bytes) to Exabyte's (10 ^ 18)
Method of storage	Centralized	Non-centralized
Structuredness	Structured	Semi-structured and not structured
Data Storage and Processing Model	Vertical model	Horizontal model
Relationship of data	Strong	Weak

Source: Wikibon

According to the updated semi-annual forecast of IDC (March 2017), in 2017, the world market for Big Data and business intelligence (BDA) will reach $ 150.8 billion, an increase of 12.4% compared to last year. It is expected that commercial purchases of equipment, software and services related to the BDA will show an average annual growth rate (in compound percentages, CAGR) of 11.9% in the period until 2020, when the revenue will exceed $ 210 billion.



Segments with the fastest growth in spending on large data and analytics: CAGR for the period 2015-2020. Source: IDC

The most significant investment in BDA technology in 2017 is expected in the banking sector, in discrete and continuous production segments, in federal / central government bodies, as well as in professional services.

In the present time, current, difficult economic conditions affect all areas, including the banking sector. In the present time, current, difficult economic conditions affect all areas, including the banking sector. Overdue debt on loans is growing and the question arises, how to manage it, how to minimize it. Very urgent is the question of obtaining the actual contact information and data on customers, for establishing contact and settling overdue debts. Such data can be both the place of residence and registration of the client, and his contact information (phone numbers, e-mail addresses), and information on his accounts and products that he uses. It is worth noting that banks have long implemented a software mechanism to establish a customer in their base, tracking the use of a product. These programs provide a large set of functions for processing and using this information. In particular, these programs are used to track the fact of the client's exit on the overdue debt on the loan or credit card. With the help of these programs, all the information about the client that is available to the bank is collected, to establish contact with the client both personal and remote via SMS, calls, e-mail. Contact with the client is supported for different purposes: to provide information on new services of the bank, to inform the client about the service that it already uses, to inform the client about the withdrawal of the overdue debt, to settle the issue of repaying overdue debts.

The last question, in the current economic conditions, is critically important, as the crediting of the population of the Republic of Azerbaijan. Today the situation of the banks of Azerbaijan is quite heavy. A number of banks stopped issuing loans in AZN, others offer only in dollars, but citizens do not want to take them. Still others do not issue loans because of an internal complicated financial situation. Debt of the population of Azerbaijan to the banks of the country continues to remain one of the most discussed problems in the country. According to

official information, in Azerbaijan, the total amount of overdue loans as of February 1, 2017 was AZN 1 633.1 million.[1]

Therefore, contact with a customer who has reached an overdue debt is important. It depends on him whether the bank employee will be able to contact the client and find ways to pay off the overdue debt, or the contact will not be established because of irrelevant information about the client. However, the speed of changing customer data is very high. As a result, the data becomes obsolete and requires constant updating.

This issue is very important from the economic point of view. Banks are forced to create additional reserves for those loans that are out of arrears. Thus, the bank freezes money that could be used directly in its operations to provide credit products [5].

Relational databases no longer allow you to store data and provide solutions for the rapid processing of large amounts of data. Here we are talking about terabytes of information. Moreover, Big Data allows you to work with both structured information and unstructured data, such as images, text, video. The former programs and hardware do not allow fully analyzing and processing such large volumes of data [6]. There is already a fairly large set of tools, technologies that allow working with Big Data, like NoSQL, Map Reduce and Hadoop [7].

These and other products are designed to work with Big Data, work on the principle of "sharing" [7]. This principle can be explained as follows. There is one database that accepts a request for certain data. This database sends this request to many other databases, and collects from them already summary information on request. Thus, the database of the requestor first deals with much smaller information and more structured information. This technology allows you to speed up the processing time of queries that require analysis of large amounts of data with different characteristics. The demand for such an analysis of data is now growing in all kinds of spheres: economy, business, science, medicine and many others. The direction on the use of Big Data in the work of the bank, in the collection of arrears is considered promising, as it is necessary to contact the debtor to collect overdue debts. To establish communication with the client, you need the actual contact information. Thanks to the Big Data technology, it becomes possible to process not only the data that the bank has to search for such information, but also to search for contact information in other places. For example: in social networks, forums and other resources, where the client could leave his current data. To process such large volumes of information, Big Data technology is required. Moreover, information in social networks can be presented in text format or even a photo or video. Consequently, the transition to work with large amounts of data is more relevant than ever in this area. To search and process customer data in social networks, databases of mobile operators, insurance companies just need Big Data. However, in this case we are not talking about the traditional view of Big Data as a statistical tool, but about the distributed Big Data.

That is, the search and analysis of data is not in one centralized database, but in several. As already mentioned, this is due to the fact that the search is conducted in several independent databases. So, there are already many examples of using Big Data for collecting information in social networks. According to CyberSecyrity.ru [8], studies of the University of British Columbia in Vancouver, using bots in social networks, showed that with the help of these bots, it is possible to get quite large amounts of information: how many shared a link, how many times a particular client Went to a certain page and much more. From this study it is clear that the search

[1] http://ru.echo.az/?p=57697

for contact information (phone, address) is possible. As it is necessary to process large amounts of data and work with them, when they are presented in various formats, it is necessary to use Big Data technologies.

A similar example of using Big Data is the search and analysis of data in social networks. The same applies to the search and analysis of data in various databases of cellular operators, insurance companies. In conditions when such databases can be legally acquired, the question of creating a database for the application of Big Data technologies is eliminated. Thus, the realization of the idea of searching for contact data of the bank's customers becomes realistic, able to bring significant contribution to increasing the speed and quality of updating such data. The presence of actual customer contacts at times improves the efficiency and speed of the collection of overdue debt, which in turn reduces the bank's costs of creating reserves. It is worth noting that similar developments already exist in a number of banks. However, they require further development, and so far these are only a few attempts to optimize the bank's work in the collection of overdue debt. Collecting and updating customer contact information has many tasks. One of them is the search and collection of contact information about customers by the bank, in order to establish contact with it in the framework of the bank's function to collect arrears.

CONCLUSION

The use of this technology can give tangible economic effect to the bank and the banking system as a whole, as more efficient collection of arrears means minimizing the costs of creating reserves for overdue loans. This area of application of Big Data technologies requires further study and development to provide a well-functioning mechanism for information support for the collection of overdue debt by banks. Moreover, further development of Big Data application in the banking sphere is possible. One of the promising directions in this area is the identification and analysis of factors that must be taken into account when granting loans. We are talking about creating a multi-variable model that would allow analyzing various parameters that affect the potential risk of non-repayment of a loan or the client's exit on arrears. Therefore, it will be possible to directly affect the quality of the loan portfolio of the bank and make it healthier.

Thus, the use of Big Data technologies to search for contact information about clients in order to realize the bank's function to collect arrears is an actual direction. The already existing practice of applying large data technologies for similar purposes creates a practical basis for the effective application of Big Data to the specified function of the bank. However, the use of such a technology is not limited to this and can be aimed at creating a multifactor model of client analysis at the stage of assessing the application for a loan. Such a model would minimize the risks of non-repayment of the loan or the exit of the client to overdue debts. In addition, these decisions would allow the bank to improve its credit portfolio and minimize economic risks.

REFERENCE

I. James Manyika's "Big data: The next frontier for innovation, competition, and productivity": James
II. Manyika, Michael Chui, Brad Brown: McKinsey Global Institute, 2011.-156 p.
III. Big Data (BigData) .http: //www.tadviser.ru/index.php/Article: Big_Data_ (Big_Data).

IV. Budzko VI High Availability Systems and Big Data. // Big Data in the National Economy 2013. P. 1619.

V. Moscow Stock Exchange "Analytical review of the BigData market"http://habrahabr.ru/company/moex/blog/256747/

VI. Provisions for possible losses on loans [Electronic resource] // Banks. RU. Electron. Dan. - M., 2015. - URL: http: //www.banki.ru/wikibank/obyazatelnyie_rezervyi_ kommercheskih_bankov / - (date of circulation: 17.09.2015).Large amounts of data [Electronic resource] // Dis. Electron. Dan. - M., 2015.URL: http://www.disgroup.ru/solutions/data_management/big_data/ - (date of circulation: 17.09.2015).

VII. Big Data Revolution: How to extract the necessary information from the "Big Data"? [Electronic resource] // StatSoft. Electron. Dan. - M., 2015. URL: http: //statsoft.ru/products/Enterprise/big-data.php - (date of circulation: 17.09.2015).

VIII. "Socio-bots" - a new threat to users of social networks? [Electronic resource] // CyberSecyrity.ru. Electron. Dan. - M., 2015. URL: http://www.cybersecurity.ru/net/136664.html - (date of circulation: 17.09.2015)