УДК 004.6

DATA ANALYSIS AND PROCESSING

N.A.Naim
Assistant at the Tashkent University
of Information Technologies
named after Muhammad al Khorazmiy,
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. The article reflects the results of research aimed at increasing the efficiency of global data processing. The proposed approach provides all the capabilities for global processing and work with data. When healthcare organizations work with Teradata, data processing increases exponentially.

Keywords: Big data, medicine, DBMS, teradata

Introduction. Many researchers are engaged in identifying features in Big data. Popularity in the study of big data appeared already at the beginning of 2010, but the greatest vector of development fell in 2020–2020. Lately, a modern person cannot imagine his day without posting photos, posts and videos on the open Internet. At the same time, machines generate even more information; smartphones and computers are to blame, and today they create more than 80% of new data. In 2018, the weight of data around the world was 33 billion TB, and by 2025 it is predicted to reach 175 billion TB. Due to the great growth of data, it needs to be analyzed and studied. It is necessary to simultaneously look at how the temperature is changing across the globe or how the decrease in water resources on the planet has occurred over the past 50 years.

Big data began to increase its characteristics as it was studied and subsequently exploited to the already familiar characteristics: huge volumes, speed of data analysis, diversity. Add authenticity and value. Anyone who can now learn to store and process large volumes of data will always have a competitive advantage in the financial market. The future is getting closer, in which big data analysis will make it possible to figure out how to distribute supplies of food, energy and medicine at the level of cities, countries, or even the entire planet. In-depth study is already helping to fight crime, and the first steps are being taken in the United States. Based on historical data on thefts, analytics department specialists predict in which areas new incidents will occur and send more patrol officers to such areas [1].

In 2019, the whole world was faced with the COVID-19 pandemic such global shocks had not occurred in more than 75 years. In the 21st century, this is the first epidemic that has so greatly affected the life cycle of people across the planet. For the first time, many countries are faced with such a large number of cases and, most importantly, mortality. Medical facilities were overcrowded and a number of organizations lacked a structured procedure for dealing with such situations. Doctors were faced with a new disease, not understanding what methods to treat, and

most importantly, save the population. To formulate a competent treatment process, it was necessary to understand the primary symptoms such as: age, gender, chronic diseases, date of illness. Those organizations that have already used Big data in their work fought the disease much faster and more successfully. Thanks to specialists from the Big Data profile, DBMSs were created that helped doctors track the course of the disease and which medicine had the best effect during therapy. Reliance on artificial intelligence has helped to significantly reduce the number of deaths, as the time for making decisions has been reduced.

Big Data in medicine could significantly expand horizons in working with the healthcare sectors and make the work of an ordinary specialist easier. At higher levels of management, it is possible to replenish specialists in those places where it is necessary, through competent distribution from unloadings to medical institutions. In the last 2 years alone, healthcare spending has increased around the world, with spending reaching more than 15% of GDP in Germany and the US. At the same time, there is no direct dependence on the increase in funds - the quality of services provided did not improve, life expectancy did not increase, and the provision of medical care was extended over time and led to critical consequences [2].

During research on Big data, it was revealed that not all DBMSs are suitable for working with medical institutions. Teradata showed itself to be the most mobile. This mobility is associated with the ability to store the largest possible amount of data. It is worth understanding that strings have a limit of 64K characters, and storing data in UNICODE requires 2 times more disk space. In turn, Taradata stores the date in the form INTEGER, the formula for converting the date to a number is as follows: (year-1900) *10000+month*100+day. It should also be taken into account that implicit conversions caused by differences in the types of fields used to perform the connection lead to an additional decrease in performance. Understanding the memory area in Teradata, the area is divided into three types, which greatly facilitates working with databases. The first type is Permanent space, which in turn includes: databases, users, tables, views, indexes, procedures, logs. Unlike other databases, Teradata does not allocate space to an object at the time of its creation. Each database or user determines the maximum possible amount of bytes that their objects can occupy - this is Maper. The entire volume of Permanent space is divided between AMPs of the system, thus each workstation has information about the space available for each database. The amount of space consumed increases as objects are created in the database. The occupied space is occupied by CurrentPern. Spool and Temporary space are temporary spaces. In turn, Spool Space is used to create files containing intermediate or result sets of data. To create files, the system uses those disk cylinders that are not occupied by permanent or temporary data. MaxSpool is the maximum possible number of bytes that the system will allocate to create user spool files. CurrentSpool is the number of bytes currently allocated to active sessions. PeakSpool - the largest number of bytes of data ever used for a transaction Temporary Space is designed to store global temporary tables; such tables are available to many users and are automatically cleared after the session.

Teradata has shown that it perfectly generates medical data that doctors can later use in their work [3].

The article reflects the results of a study aimed at increasing the efficiency of global data processing. The proposed approach provides a whole range of possibilities for global processing and working with data. When implementing the work of medical organizations on Teradata, information processing will increase by 20%. It is necessary to increase research in the field of semi-structured databases and expand standards when working with microdata; all this research carries enormous potential for implementation not only in the Russian market, but also in the global one. Research will help circumvent sanctions imposed on the financial sector of the

economy. These solutions are used for leading financial organizations that are interested in reducing technological costs and introducing advanced technologies in their work.

References

[1] N.A. Naim., Big data in modern business, Сборник научных статей VIII международной научно-практической конференции. «Big data и анализ высокого уровня». Минск – 2022. – С. 221-223.

[2] N.A. Naim., Modern trends in bigdata and datascience, Сборник научных статей VIII международной научно-практической конференции. «Big data и анализ высокого уровня». Минск– 2023. – С. 62-63.

[3] N.A. Naim., R.R. Qodirov., Application of big data technology, Сборник научных статей VIII международной научно-практической конференции. «Big data и анализ высокого уровня». Минск– 2023. – С. 79-82.

author's contribution

Naim Nodira Abdujalolovna – development of the matter, setting of problems, formation of the article.

АНАЛИЗ И ОБРАБОТКА ДАННЫХ

Н.А. Наим
Ассистент в ТУИТ имени
Мухаммада ал-Хорезми

Аннотация. В статье отражены результаты исследований, направленных на повышение эффективности глобальной обработки данных. Предлагаемый подход предоставляет все возможности глобальной обработки и работы с данными. Когда организации здравоохранения работают с *Teradata*, обработка данных увеличивается в геометрической прогрессии.

Ключевые слова.: *Big data*, медицина, СУБД, *teradata*