УДК 519.257

BIG DATA TECHNOLOGIES IN TECHNOLOGICAL AND BUSINESS PROCESSES AUTOMATIZATION

M. Y. Abdul-Azalova
Senior lecturer of the Information technologies Department of the TUIT, Postgraduate student of the TUIT

N. M. Mamatova
Deputy Dean of the joint faculty of the JFIT TUIT-BSUIR, Postgraduate student of the TUIT

Joint faculty of JFIT TUIT-BSUIR, Republic of Uzbekistan.
Tashkent University of Information Technologies named after Muhammad al-Khwarizmi, Republic of Uzbekistan.
E-mail: bonu444@mail.ru.

M. Y. Abdul-Azalova
She graduated from the Tashkent University of Information Technologies named after Muhammad al-Khwarizmi. Senior lecturer of the Department of «Information technologies» of TUIT. Postgraduate student of the TUIT. Conducts research on algorithms and adaptive models of an extensive multiparametric e-business management system.

N. M. Mamatova
She graduated from the Tashkent University of Information Technologies named after Muhammad al-Khwarizmi. Deputy Dean of the joint faculty of SFIT TUIT-BSUIR. Postgraduate student of the TUIT. Conducts research on algorithms and models of integrated, intelligent information exchange systems.

Abstract. Big data is a collection of huge useful information that cannot be read using standard computing structures. Big data is not just data, it has already become a whole area that includes a set of tools, contexts, and structures. It uses complex data sets to select direction, course, and direct management within organizations. This article examines the process of automating technological and business processes using Big Data technologies, identifies the main stages of implementation of Big Data technologies in the automation of technological and business processes, and analyzes the experience of implementing Big Data technologies in modern industrial enterprises.

Keywords: business-process, advanced analytics, ARIS, DIRECTUM, web analytics, predictive analytics, personalization, dynamic analysis, sentiment analysis, Big Data, e-commerce, e-business.

Introduction.

The transition to «digital manufacturing» and technological development based on digital technology is the basis for the global competitiveness of modern enterprises. Previously, the term «digital manufacturing» was understood as a set of application systems that were mainly used at the stage of technological production, namely, to automate program development processes for CNC machines, to automate process development for assembly, to automate tasks related to job scheduling in robot programming, and to integrate with shop floor systems (or MES, Manufacturing Execution System) and ERP resource management systems. In recent years, with the advent of new breakthrough technologies, the term has been interpreted more broadly. Today, «digital manufacturing» is primarily understood as the use of digital modelling and design technologies for both products and products, as well as production processes throughout the entire lifecycle. In essence, we are talking about creating digital twins of the product and its production processes.

This article examines the process of automating technological and business processes using Big

Data technologies, identifies the main stages of implementation of Big Data technologies in the automation of technological and business processes, and analyzes the experience of implementing Big Data technologies in modern industrial enterprises.

Materials and methods.

Experts believe that the changes in modern industry (some of which are already taking place) that «digital manufacturing»implies will take place in the following key areas:

– Digital modelling – the concept of a digital twin, i.e. manufacturing a product in a virtual model that includes equipment, production process and plant personnel, is developing.

– «Big Data» (Big Data) and business intelligence that emerges from the manufacturing process.

– Autonomous robots, which will gain greater industrial functionality, independence, flexibility and execution ability than the previous generation.

– Horizontal and vertical systems integration – most of the huge number of information systems currently in use are integrated, but there needs to be greater collaboration at different levels within the enterprise as well as between different businesses.

– The Industrial Internet of Things, where information from a large number of sensors and equipment coming from production is networked together.

One of the hallmarks of «digital manufacturing» is the presence of an intelligent control system, i.e. the ability to tightly integrate existing process equipment and obtain a wide range of process information from anywhere in the production ecosystem [1].

Today, data is one of the most important components of society and every person's life. The modern stage of society is characterised by a constant increase in the volume of data. Data comes from many different sources, such as data from GPS navigators, satellites, Internet queries, social networks, and data from the IoT (Internet of Things). The structure and composition of this data is often not defined. Big Data (Big Data) has the following properties: huge size, heterogeneity and disorderliness, require fast processing. Big Data technologies are a set of tools, approaches and methods for processing both structured and unstructured data of huge size for further use.

The main Big Data technologies and tools include:

– Hadoop & MapReduce;

– NoSQL databases;

– Advanced analytics (statistics, predictive analytics and Data Mining, linguistic text processing);

– Data Discovery class tools.

Practical implementation of Big Data technologies are modern neural networks and derivative systems based on them, such as pattern recognition systems, simulation modelling, machine learning and predictive analytics. Big Data technologies are widespread in the banking, telecommunications, industry, healthcare, energy, insurance and trade sectors. Large industry has been collecting huge amounts of data for many years to improve product quality and production efficiency [2]. The main materials for research on the subject area are: a sample of scientific and professional works of domestic authors in the field of Big Data technologies, processing of huge amounts of data, Internet resources on the subject under study. The research methods are comparative and system analysis, logical approach.

Results.

Technologically, the operation of a highly automated (including extensive use of industrial robotics) digital enterprise is summarised as follows. Using IoT technologies, vast amounts of information are collected in physical space and sent into cyberspace, where it is analysed using artificial intelligence. The results of this analysis are fed back into the physical dimension, where management decisions are made on the basis of this information.

Big Data technologies make it possible to automate technological and business processes, which leads to an increase in the speed of business reactions to external and internal disturbances. The economic effect is achieved by increasing the transparency of processes, improving the quality of planning, introducing pervasive management by deviation (or target), increasing the speed of identifying the causes of deviations, constant standardisation and standardisation of best practices. Implementation of Big Data has a very specific goal – to implement a dynamic business management model that enables rapid business

reactions to external and internal perturbations. If perturbations occur in the system, be it instability in sales, variations in production, deviations in supply, the adaptive business management model allows the business to quickly «readjust» to the best strategy, which ensures it dramatically improves competitiveness. Moreover, the adaptive business management model uses instabilities, variations and deviations to continuously improve the structure and parameters of the business [3].

The general scheme of application of Big Data technologies in the process of automation of technological and business processes in the enterprise is shown in the figure 1.

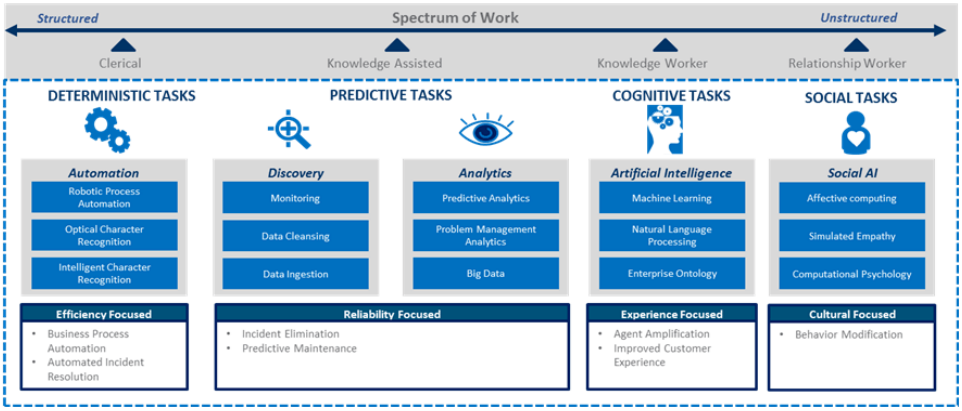

Figure 1. Big Data technologies in process and business process automation

Manufacturers install sensors on key pieces of equipment to collect information in real time. The collected and processed data is sent to all divisions of the company to enable interaction between structural divisions and to make appropriate management decisions.

This information can be used to improve service (prevention of downtime, equipment breakdowns), to create targeted marketing offers [4].

Continuous monitoring of key indicators makes it possible to identify the problem and take the necessary measures to solve it. Modern systems make it possible to monitor the technological process and identify influencing factors using any Web-browser. Such solutions make it possible to turn production data into the information required for effective enterprise management [2].

The main stages of the implementation of Big Data technologies in the automation of technological and business processes are as follows.

1. Data extraction from sources (most of the data is generated at the level of sensors, APCS, MES, ERP).
2. Data storage (appropriate data storage independent of production servers.
3. Data processing (data normalization so that later it is possible to apply the tools of «agile analytics», statistical control and analysis, numerical modeling).
4. Data analysis. «Flexible analytics» makes it possible, with the help of wizards, to organise client-oriented analytics for themselves simply and quickly. That is, any trained worker sets up for himself the reporting, which he needs to work: reports, graphs, histograms, regressions, Shuhart charts. This dramatically increases the efficiency of data analysis in business.
5. Digital modeling allows you to find «digital twins» not only of technological objects, but also of business processes. When a specialist has a «digital twin», he can quickly find the best rules, technological modes, procedures, regulations. Besides, «digital twins» allow to realize operative end-to-end planning of a stream of creation of value according to strategic purposes that provides both operative synchronization of objects of a stream, and their optimization.
6. Optimization.
7. Forecasting.

Big Data technologies, as a key component of the Industrial Internet of Things, are already widely used in many enterprises, enabling workers to improve their existing skills and the enterprise to function more efficiently.

Here are a few examples of the use of Big Data technologies in industrial enterprises.

ThyssenKrupp AG, one of the world's leading lift manufacturers, has implemented MAX, a system that collects data from sensors installed in the company's lifts via the Internet of Things, and from that builds models on the Azure Machine Learning platform that can prevent incidents before they occur and communicate a specific breakdown code to technicians to reduce maintenance time.

General Electric Oil & Gas, a manufacturer of high-tech equipment for the oil and gas sector, uses Big Data to minimise the time of unplanned oil production shutdowns.

Using historical sales data and optimization algorithms, Nestle is using big data analytics to optimize the production chain, reducing errors in forecasting material demand by half, as well as losses from holding excess working capital, delays in production, etc.

Each computer component manufactured by Intel must pass approximately 19 000 different tests before it is released on the market. The use of Big Data technologies makes it possible to leave only a portion of these tests based on data analysis of the entire production process, consequently reducing component testing time and costs.

Another important issue is the design and control of business processes.

The Business Process Management (BPM) cycle consists of a series of sequential steps which are managed by different employees within the organisation and supported by different applications. In this article, we will try to explain how this cycle is performed using the popular tools offered by DIRECTUM and IDS Scheer (Figure 2).

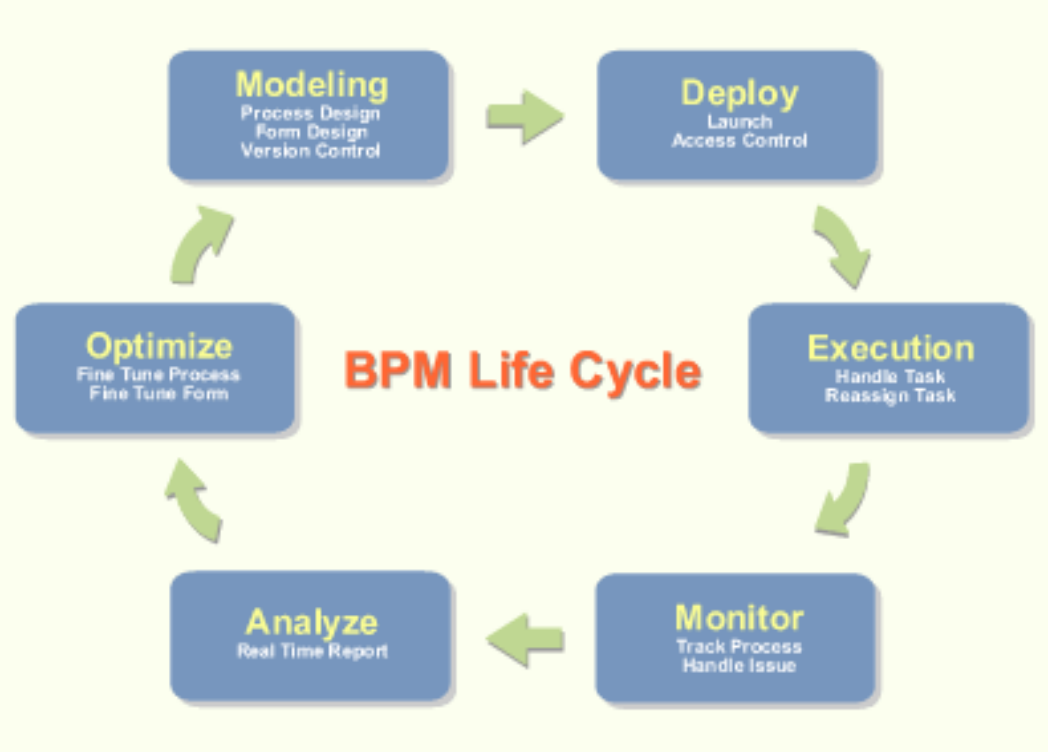

Figure 2. The business process management cycle

In the strategic business process definition phase the most general goals of the organization are defined and the scope of its activities is outlined. The managers of the company formulate the key tasks and define the upper level processes; they also allocate areas of responsibility for the processes. This phase is supported by business process design systems. In particular, in the ARIS family of IDS Scheer for such tasks there are ARIS BSC Solution, ARIS Business Optimizer and ARIS Toolset products. The latter provides simulation of the company's strategic objectives, value chains and top-level processes.

The process design stage involves the creation of «as is» and «as should be» models. Business analyst (consultant, «modeler») at this stage describes the processes with their full detail. In the ARIS

product family, the design tools include ARIS Business Architect, ARIS Business Designer and ARIS Toolset (or its lighter version ARIS Easy Designer).

The created descriptions are then coordinated with the process owners, their immediate or potential participants and department managers. At this stage, the DIRECTUM content management environment plays the role of a system where the descriptions are agreed upon and subsequently the information about the processes is published.

Now that the business processes have been designed, they must be implemented. Here, the supporting software tools are required not only to plan the implementation process, but also the actual transfer of models and related dictionaries (e.g, underlying organizational charts) in those or other information systems designed to further manage the execution of business processes. ARIS family contains tools for the transfer of models, specialized tools to manage the implementation of processes through various BPM-systems, modules and SOA-boxes, as well as utilities to display the models in software products under development (ARIS for SAP NetWeaver, ARIS UML Designer). In turn, DIRECTUM supports various options for the import and generation of schemas and dictionaries, as well as providing documentation support for business processes.

The designed processes are executed by different people and with the help of different information systems. The role of a «conductor» is played here by either a separate BPM-tool, or an SOA-bus (depending on process properties, purpose and level of automation), or software such as BPM/workflow-modules of enterprise content management (ECM). In the ECM system DIRECTUM in particular this is the module «Business Process Management».

Execution of business processes in automated systems simplifies the collection of related data and allows for analysis of processes and their further optimization. Analysis and optimisation are the key tasks of the controlling phase. Corporate IS (ERP, CRM, ECM) provides statistical information about the course of processes, about their individual stages, about the passage of certain control points. The controlling module (in ARIS it is Process Performance Manager) processes this information, comparing it with a process model. As a result, the business analyst or manager is notified when the process is not running smoothly, identifies bottlenecks, finds the busiest performers and evaluates other performance indicators.

In addition, at all stages of the business process management cycle the task of change management must be addressed from time to time: models and their mapping to the BPM system, BPM/workflow modules and other IS elements must be kept up-to-date.

Two key challenges arise during process implementation. First, in order to accelerate process descriptions, reduce the amount of routine work, unify the benchmark processes and their execution, you need to automatically synchronize the designed models with the systems that automate the execution of business processes. In many cases these systems are ECM solutions and BPM/workflow modules included in them. Secondly, it is often necessary to synchronize information about static entities such as enterprise structure, technical resources diagram, top level goals and functions diagram, network topology, etc. with business applications. This information is sometimes ancillary, but often meaningful on its own and can be actively used in various information systems.

Process diagrams can be transferred from the description environment (ARIS) to the execution environment (DIRECTUM) as well as in the opposite direction (Figure 3). In ARIS the processes are described in the form of so-called eEPC diagrams while in DIRECTUM in the Business Process Management module free and hard generic routes are used for this purpose. When transferring blocks of type routes, certain rules are applied to match them with the elements of eEPC diagrams. It is important to understand that the transfer of eEPC diagrams cannot.

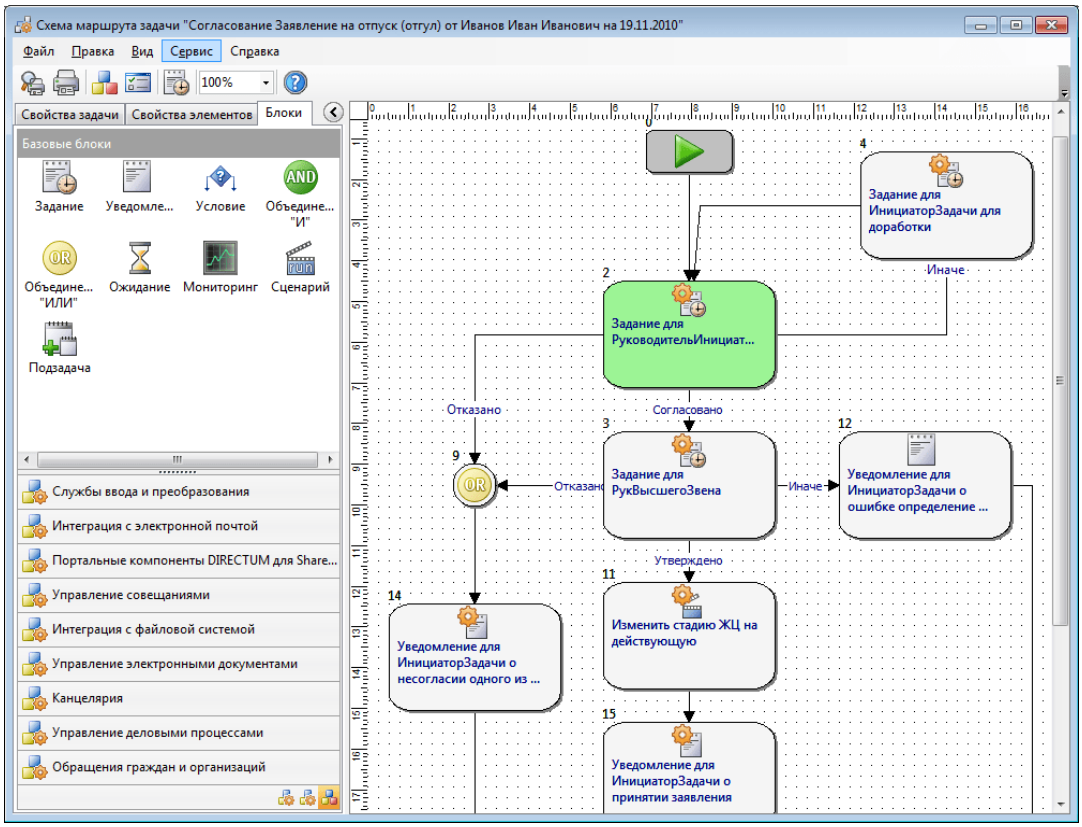

Figure 3. Example of transferring the correct model from ARIS to DIRECTUM

The layout of the typical route only requires a little manual adjustment of the block coordinates.

When transferring schematics from ARIS to DIRECTUM it is also necessary to consider several limitations. Firstly, in most cases end-to-end processes are designed in ARIS, not only in ECM system but also in other corporate applications (ERP, CRM etc.). In other words, DIRECTUM is likely to automate only a part of the process. This limitation can be circumvented by fixing in the model (e.g. with attributes) those process functions that should be reflected in the DIRECTUM typical routes. Secondly, a process transferred to DIRECTUM cannot be started immediately. In any case, it is necessary to fine-tune it: to select and set route parameters, to precisely match the roles, to set restrictions on the execution of blocks (for example, the presence of an attached document in the task) or to simply adjust the coordinates of blocks for a more visual representation of the route. Nevertheless, automated model synchronisation can significantly reduce the share of routine work.

The third limitation is related to the fact that in real life, modellers do not always thoroughly observe the requirements of notation (process description language). Often the models they create lack the necessary links, use objects similar in appearance but different in meaning, use different instances of the same object instead of instances of different objects, lack mandatory information (for example, the performer is not specified), etc.

The reverse transfer of models from DIRECTUM in ARIS is less prone to error, it is easier because, as an automated process, initially described more strictly. Such a transfer is necessary when ARIS is implemented later than DIRECTUM. It allows to fix in system modeling those developments in the field of regulation of business processes which have been already made by the moment of implementation of ARIS. Nevertheless more typical situation when ARIS contains reference models which find their reflection in DIRECTUM. When transferring eEPC diagrams from ARIS to DIRECTUM and back again, AML (ARIS Modeling Language), the XML-based ARIS model description language, is used.

In addition to process diagrams, static diagrams developed in ARIS can also be transferred. An example is the organization chart, the data from which are migrated to the directories of departments and employees supported in DIRECTUM and further used in the standard routes.

The transfer of organisational diagrams, like the transfer of process diagrams, follows certain rules. The transfer script «understands» to which type this or that block belongs and what are the relations between the blocks. For example, the unit object in an ARIS diagram is converted into a directory entry, the information of the associated object «position» enters into the «position of the head» of the unit, and the information about the person holding the position is transferred to the directory of employees and linked to the directory entry of units (Figure 4).

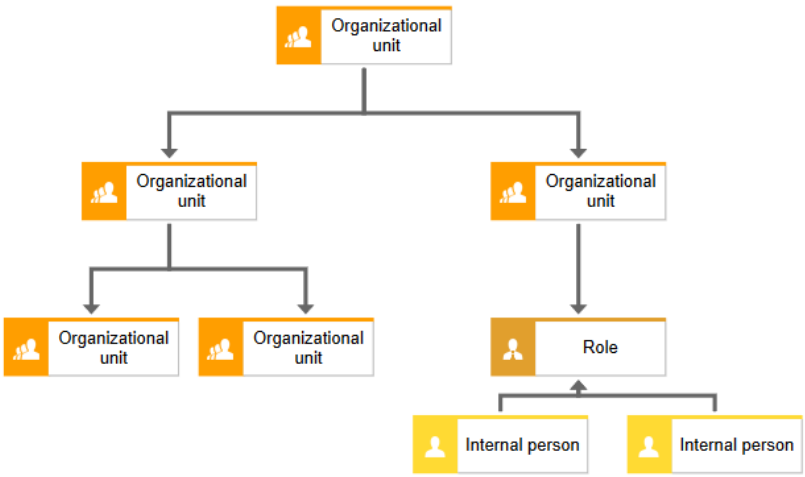

Figure 4. Example of transferring the organisational structure from ARIS Toolset to DIRECTUM

The reverse generation of an org chart is also possible using the contents of the DIRECTUM department directory.

The result of the design in ARIS Toolset are not only process schemes. Like most similar process design and support systems, ARIS Toolset generates various reports and documents. These documents can be useful to a small group of people, such as consultants, analysts and managers (eg process ownership matrixes, lists of processes and their diagrams, staff schedules), as well as wide range of employees (work instructions, lists of employees in various departments, etc.). DIRECTUM can store documents generated in this way, help to coordinate them and bring them to all interested employees of the organization.

The essence of controlling is in the regular assessment of the degree of achievement of set objectives, continuous improvement and optimisation of processes. When we talk about quality of building a system of business processes, we mean achievement of the highest possible speed of their execution, simplicity and transparency, i.e. absence of unnecessary steps and bottlenecks in processes. Therefore, the tasks of controlling are, in particular, the continuous collection of statistical information on the execution of business processes and verification of compliance of actual parameters with predetermined standards.

In ARIS product family these functions are performed by ARIS Process Performance Manager (ARIS PPM). It forms base of statistics, in it the control indicators are set, and by means of it stakeholders are informed on those or other revealed problems of actual execution of processes. ARIS PPM is focused on analysis of end-to-end business processes, which involve various corporate business applications – financial and production, budgeting, customer relationship management and of course ECM systems. Each of them can either have its own workflow/BPM module, or implement the process execution within itself in some special way. The concept of a controlling system is therefore based on the analysis of how the process passes through control points. Information about the fact of passing such a point must be generated by the business applications involved in the process execution and delivered to the controlling system.

The task of DIRECTUM, working in conjunction with ARIS PPM, is to provide information about the start and completion of the job, about his passage through a particular stage of the route model, on the composition of the route parameters, as well as – if necessary – information about the execution of certain

scenarios, access to documents and directories.

Once synchronised, the process and the transferred dictionary data will subsequently change – and therefore the model transfer has to be repeated. The model and its reflection must be identical: this is the essence of the change management task. As in the case of the model synchronization discussed above, two directions can be distinguished here: The search for the ECM system objects affected during the model modification and the search for models which might have become obsolete due to a change of their implementation in DIRECTUM. In the case of DIRECTUM and ARIS the result of such a search is a discrepancy report. Manually generated or scheduled, these reports allow to compare the structures and processes of the DIRECTUM execution environment with the ARIS models.

The above described integration of DIRECTUM and ARIS tools enables faster implementation of the BPM system, timely updates of processes and structures in the DIRECTUM environment and simplified controlling of business processes.

The use of Big Data also has a number of problems. The main one is the cost of data processing, which includes expensive equipment and the cost of salaries for qualified professionals who are able to service huge amounts of information. The second problem is bias. If a study produces not 2-3 but numerous results, it is very difficult to remain objective and isolate from the total flow of data only those that will have a real impact on the state of any given phenomenon. The third problem is the protection of Big Data. Methodologies for protecting information systems in the classical three-tier architecture are not applicable to new technologies. There is a need to create and train a new class of Big Data security specialists [7].

Conclusion. Big Data technologies are now quite a workable set of technologies used in almost all areas of human activity and have great potential for further development.

Unfortunately, according to experts, Uzbekistan is still about 3-5 years behind the leading countries of the world in the use of Big Data technologies. The reasons for this lag are the low level of automation, the scattered nature of the data being collected, and the insufficient number of real projects. Technologies used in the West cannot always be adapted to the Uzbek reality. In addition, there is an obvious shortage of Big Data specialists.

The introduction of Big Data technology requires not only technical support, but also organisational support. The first involves organizing data extraction, data storage, unified workstations for analysis, digital modeling, optimization and forecasting. The second direction will require the formation of appropriate qualifications in the Big Data business. Professionals with the new qualifications of «data engineers», «data scientist» for modelling, optimisation and forecasting are needed. In addition, training of Big Data technologists, planners and managers from business will be required.

Nevertheless, the potential of the uzbek big data market is enormous and in the coming years its rate of development will be many times higher than that of the global market.

References

[1] H. Liu and X. Wang, Shanghai,2018, «Study on the application of big data in accurate marketing of cross-border e-commerce in China,» 2018 IEEE 3rd International Conference on Big Data Analysis (ICBDA), pp. 24-27.

[2] X. Zhao, «A Study on the Applications of Big Data in Cross-Border E-Commerce,» 2018 IEEE 15th International Conference on e-Business Engineering (ICEBE), Xi'an, 2018, pp. 280-284.

[3] E. Makki and L. Chang, «Leveraging social big data for performance evaluation of E-commerce websites,» 2016 IEEE International Conference on Big Data (Big Data), Washington, DC, 2016, pp. 2525-2534.

[4] Shahriar Akter & Samuel Fosso Wamba,2016, «Big data analytics in E-commerce: a systematic review and agenda for future research», IIM University of St. Gallen.

[5] S. Suguna, M. Vithya and J. I. C. Eunaicy, «Big data analysis in e-commerce system using HadoopMapReduce,» 2016 International Conference on Inventive Computation Technologies (ICICT), Coimbatore, 2016, pp. 1-6.

[6] S. Pradeep and J. S. Kallimani, «A survey on various challenges and aspects in handling big data,» 2017 International Conference on Electrical, Electronics, Communication, Computer, and Optimization Techniques (ICEECCOT), Mysuru, 2017, pp. 1-5.

[7] H. Xu, K. Li and G. Fan, «Novel Model of E-Commerce Marketing Based on Big Data Analysis and Processing» 2017 International Conference on Computer Network, Electronic and Automation (ICCNEA), Xi'an, 2017, pp. 80-84.

ОПТИМИЗАЦИЯ БИЗНЕС ПРОЦЕССОВ ПОСРЕДСТВОМ BIG DATA

М. Я. АБДУЛ-АЗАЛОВА
Старший преподаватель кафедры «Информационные технологии» ТУИТ, соискатель ТУИТ

Н.М. МАМАТОВА
Заместитель декана совместного факультета СФИТ ТУИТ-БГУИР, соискатель ТУИТ

Совместный факультет СФИТ ТУИТ-БГУИР, Республика Узбекистан
Ташкентский университет информационных технологий имени Мухаммада ал-Хорезми, Республика Узбекистан
E-mail: bonu444@mail.ru

Аннотация. Большие данные - это совокупность огромной полезной информации, которая не может быть прочитана с помощью стандартных вычислительных структур. Большие данные - это не просто данные, они уже стали целой областью, включающей в себя набор инструментов, контекстов и структур. Они используют сложные наборы данных для выбора направления, курса и непосредственного управления внутри организаций. В данной статье рассматривается процесс автоматизации технологических и бизнес-процессов с использованием технологий Big Data, определяются основные этапы внедрения технологий Big Data в автоматизацию технологических и бизнес-процессов, анализируется опыт внедрения технологий Big Data на современных промышленных предприятиях.

Ключевые слова: бизнес процессы, продвинутая аналитика, ARIS, DIRECTUM, веб-аналитика, прогнозируемая аналитика, персонализация, динамический анализ, анализ тональности, большие данные, электронная торговля, электронный бизнес.