

**D.K. Ilyassov<sup>1\*</sup>, G.A. Raikhanova<sup>2</sup>**

<sup>1</sup>*Narxoz University, Almaty, Kazakhstan*

<sup>2</sup>*Karagandy University of the name of academician E.A.Buketov, Kazakhstan*

<sup>1</sup>*didar.ilyassov@narxoz.kz, <sup>2</sup>gulnurraihanova@mail.ru*

<sup>1</sup><https://orcid.org/0000-0001-6150-6492>, <sup>2</sup><https://orcid.org/0000-0001-5355-4590>

<sup>2</sup>*Scopus Author ID: 57197807219,*

<sup>1</sup>*Researcher ID: AHC-1814-2022*

### **The challenges in using economic big data analytics in higher education**

#### **Abstract**

*Object:* To identify key challenges in using Big Data technologies in public administration of education.

*Methods:* A grouping method of research OLTC (opportunities, limitations, trends and challenges), survey, structural equation modeling, confirmatory factor analysis.

*Results:* This study highlights the importance of big data analytics adaptation for organizational performance mediated by knowledge management. The theoretical concepts of Big Data Analytics are building up with new challenges. In particular, Big Data Analytics application in education has its own further development issues addressed for decision-making and problem solving of public administration.

*Conclusions:* This study outlines the priority directions how to apply Big Data Analytics in education administration at higher education institutions (HEI) in Kazakhstan.

**Keywords:** Big Data, Big Data Analytics, prediction analytics, student performance management.

#### **Introduction**

Nowadays, the data as resource should be analyzed effectively and rapidly. Even the sources of information itself are becoming so broad and big, bringing people to the term of Big Data. The 'Big Data' term emerged in mid-1990s, meaning large, complex and changing collections of information. 'Big Data' mostly exceeds the conventional capacity of traditional databases of organizations (Weiss and Indurkha, 1998). During past many decades the power of organizations are increasing hundreds times leading to increase of data as well. It was investigated that by 2020, around 1.7 megabytes of new data will be created each second for every human on the Earth. After that, our universe will have accumulated in it 44 trillion gigabytes of data (McAfee and Brynjolfsson, 2012).

Education system around the world is facing with big data and data analytics as a new field to study. In particular, twofold use of data analytics might be highlighted as decision making for education administrators and teaching for teachers.

Service sector in Kazakhstan is competitive comparatively with other sectors of economy. Meanwhile Kazakhstani Education System might be highlighted as one of the rapidly developing service industries. Education institutions are dealing with different problems, as well as with the question how to improve their services to meet customer needs.

This study highlights the importance of big data analytics adaptation for organizational performance mediated by knowledge management.

#### **Materials and Methods**

This research is mostly focused on Big Data challenges and opportunities. More specifically, Programs in R and Python are used as the most popular Data Science languages. These Models in Data Science solve a wide range of problems as following: Human resource planning, client retention, improving performance indicators, etc. Moreover, Big Data Analytics in Education plays a crucial role in sustaining future growth in terms of uncertainty and changes. Today, automated activities everywhere — camera or scan usage, satellite TVs, distance education, robot usage are an evidence of New Trends.

Big Data is increasing tremendously during recent years. Powerful computers, mobile phones, telecommunication systems make them to rise day by day. Enterprise Data, VoIP, social media, sensors, and devices

\*Corresponding author's e-mail: [didar.ilyassov@narxoz.kz](mailto:didar.ilyassov@narxoz.kz)

increase Big Data by above 50 times. It is obvious being lost among them. Figure 1 illustrates the trends of knowledge gap and execution gap increase. The increase of available customer data overtakes analytic and execution capabilities with the big lags. Hence, the widespread use of Big Data technologies is unavoidable.

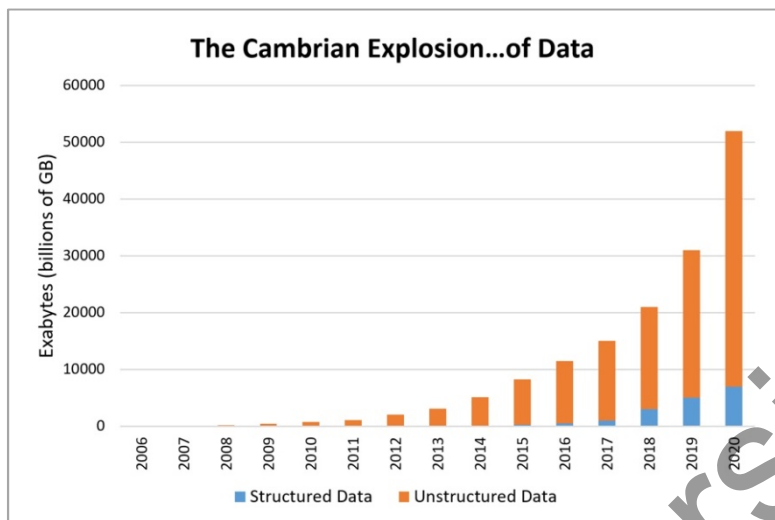


Figure 1. Knowledge and Execution Gaps

Note – Taken from the source (Cprime An Alten Company, n.d.)

The prior decade is described by launching of new digital technologies reshaped the concept of Big Data Analytics. The global social network emergence leads to the new network-related prediction models and targeting strategies on the consumer needs.

The detailed consumer data is going to rise. IT People even may derive the actionable insights by using various types of tools. Big Data analytics are becoming necessity for enterprises adding more value and smart insights. From this point of view, Data Analytics is building decision-making models added precious value by information employment.

New Big Data is also increasing for monitoring the educational systems inspiring to use new technologies in the design and delivery of public services. Due to this context, looking forward for 10–30 years, the policymaking and monitoring of education is going to make a tradeoff between rationality and the risk of creativity in using a big amount of data.

Overall, the theoretical framework of Big Data Analytics is still building facing with new challenges and limitations. Thus, Big Analytics is still unknown or unclear. Nevertheless, some significant attempts to adopt Big Data Analytics in education are already done. The combining of such studies are challenging and adds a great value as a scientific value for this study.

### Literature Review

For many years, researchers in education system used a small amount of data due to limited tools of gathering and analyzing data. The commonly used surveys and interviews have a limited interpretative power and latency validity (Daniel, 2017).

The new technologies of Big Data Analytics supply possibilities for students to deal with transactional data. The further development of Big Data Analytics enhances the relationship with strategic decision-making.

The problem of working with big data in education is obvious. The following questions, as in adapting Big Data Analytics, remain in front of employees: Why, How, What, How to evaluate. The simple answer to them might be having comparable datasets in the world. These datasets can be standardized, have common methods, tools, traits. Because of this trend, the theoretical framework of Big Data Analytics continues to develop (Lane, 2019).

In combination, big data and innovative teaching process may be applied in improving making decisions, sharing insights, discovering new things, and optimizing and simplifying the processes of learning. It is possible to implement thematic analysis to build a model for higher institutions to control their plan on big data application (Miftachul et al., 2016).

The role and value of Big Data analytics remain in Phases of Big Data Analysis: descriptive, predictive, and prescriptive (Tulasi, 2013). Similarly adhering phases of Big Data Analysis, Mohsen et al. (2018) of-

ferred a special model for the concept of using analytics in higher education. In this study, student life cycle and analytics category are related as follows: Pre-student — descriptive, predictive. Student — predictive, prescriptive. Post student — prescriptive, descriptive. The main issue that Big Data Analytics may face is student retention in higher education institutions.

During Big Data Analytics application, the next key factors are important to consider: 1) Vision and Plan; 2) Scalability; 3) User-friendly interfaces; 4) Up-to-date; 5) Real-time collaboration; 6) Quick installation, maintenance, and upgrade; 7) Reliability and security. These technical issues have to be under concern of education institutions (Mohsen et al., 2018).

To teach Big Data Analytics at universities, some studies develop its curricula and programs. Ines Mergel (2016) suggests 13 modules for a big data syllabus that extend Mason's PAPA model of ethical considerations for the information age. According to Mason's PAPA model, there are four types of data dimensions: privacy, accuracy, property, accessibility. The author also adds that the contemporary dimensions of data are the ethical, technological, process, organizational and institutional change, and analytical dimensions.

As an importance of Big Data Analytics for education administration and modernization answers on "How advances in big data are likely to transform the context and methodology of monitoring educational systems within a long-term perspective (10–30 years) students' privacy, educational equity and efficiency, student tracking, assessment and skills?" (Berendt et al., 2017). These questions are considered for education modernization.

Big Data is a new phenomenon in education; its conceptual relevance, challenges and limitations are unknown (Daniel, 2017). Therefore, some studies argue against Big Data in education. The opportunities, limitations, trends and challenges (OLTC) of Big Data Analytics are grouped as follows:

**Opportunities:**

- predicted academic success of students (Mergel, 2016);
- prognostic analytics can be used in digital courses (Mohsen et al., 2018);
- amount of information in education (Daniel, 2017);
- advanced clustering techniques (McNeely & Hahm, 2014).

**Limitations:**

- privacy issues (Mergel, 2016);
- limited to correlation and predictive models (Daniel, 2017).

**Trends:**

- digital Human Planning (Ashraf, 2020).

**Challenges:**

– forecasting, decision-making, resource allocation, performance (Mohsen et al., 2018). To conclude, the challenges of Big Data analytics lead to necessity of building a conceptual framework of Big Data Analytics in education.

The impact of BDA application to organizational performance is based on resource-based theory which stands for organizational performance dependence on contributed resources. In addition, BDA provides competitive advantages to enhance organizational performance. To improve organizational performance companies increase knowledge through BDA adoption. However, some studies state a partial mediating of knowledge management between BDA adoption and organizational performance (Shabbir & Gardezi, 2020). Figure 2 provides such relationship stated for HEI in Latin America (Sekli & Vega, 2021).

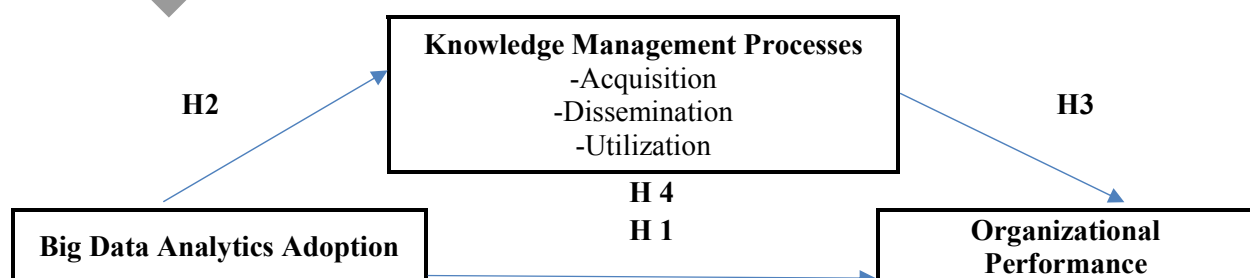


Figure 2. The interdependence of variables based on theoretical framework (Sekli & Vega, 2021)

Note – Compiled by authors on the basis of (Sekli & Vega, 2021)

It is reasonable to test the importance of BDA adoption and knowledge management for Kazakhstani HEIs' performance.

### Results

65 members of Narxoz University were surveyed to test these hypotheses. The structural equation modeling as the data analysis was applied. Table 1 presents the confirmatory factor analysis results. All four hypotheses are accepted.

Table 1. Results of hypotheses tests

Hypothesis	Standard error	p-Value	Result
H1	0.161	0.009	Accepted
H2	0.261	0.003	Accepted
H3	0.051	0.0005	Accepted
H4	0.365	0.001	Accepted

*Note – Compiled by authors*

All accepted hypotheses highlight the importance of big data analytics adaptation for organizational performance mediated by knowledge management.

### Discussion

The modernization of public administration by the use of Big Data new technologies is broad. Today, the importance of Learning Analytics is growing up. Learning Analytics improves student success and retention to educational institution. Examples of that might be massively open online courses (MOOCs), dashboard applications, predictive modelling, E-learning, etc.

One of the models of Learning Analytics is the Rasch model. It is used as the framework to analyze and somehow differentiate the students' performance and abilities with their response to a test. This whole paradigm is named as IRT, which stands for Item Response Theory (Zachary, 2017).

The important feature while analyzing Big Data is to use framework that utilizes the principle of Six Sigma (SS) at one of the public universities in the US. This framework was established and organized to improve the success of every particular student (Chad et al., 2017). Also, the concept of DMAIC (define, measure, analyze, improve and control) methodology is used for incorporation of the principles of Big Data.

Open Data and Big Data are intensively used in the education processes applying new methodology and tools considering the conventional challenging issues and proposing innovative solutions, as well as new prospective solutions in the education (Drigas & Leliopoulos, 2014).

In general, the variety and wide application borders of Big Data lead to many dimensions. Firstly, it is started from 3 Vs (volume, velocity, variety). The range of applicability of Big Data is something that can be developed in any direction and the importance of this issue is another challenging task that should be tackled. Figure 3 depicts the 3 Vs of Big Data as the early attempt of measuring Big Data.

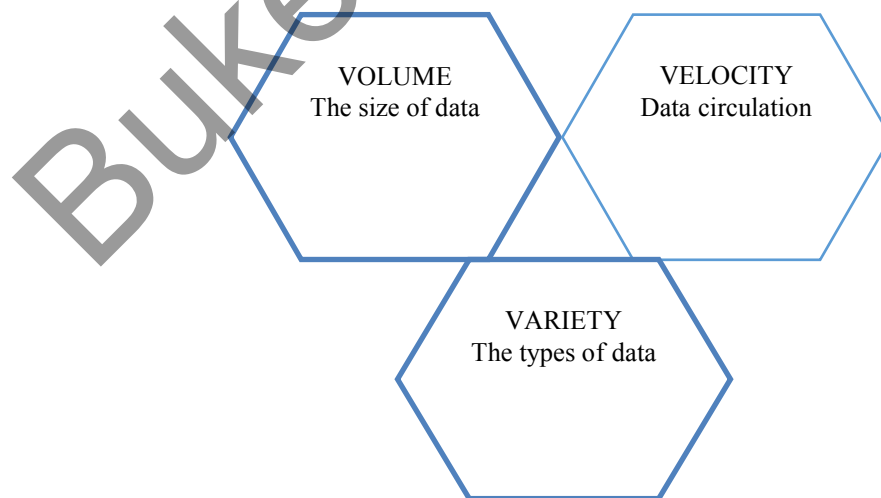


Figure 3. 3 Vs of Big Data

*Note – Compiled by authors on the basis of (Daniel, 2017)*

To understand the concepts of the 6 Vs, Ben Daniel (2017) reflects in his work the following terms: Volume, Velocity, Veracity, Variety, Validity, Value (Figure 4).

In 6 Vs conceptualization of Big Data characteristics, Validity and Value were added as additional dimensions in Figure 4. Validity shows rightness of the used data. Value means statistical, hypothetical, correlation and modelling significances.

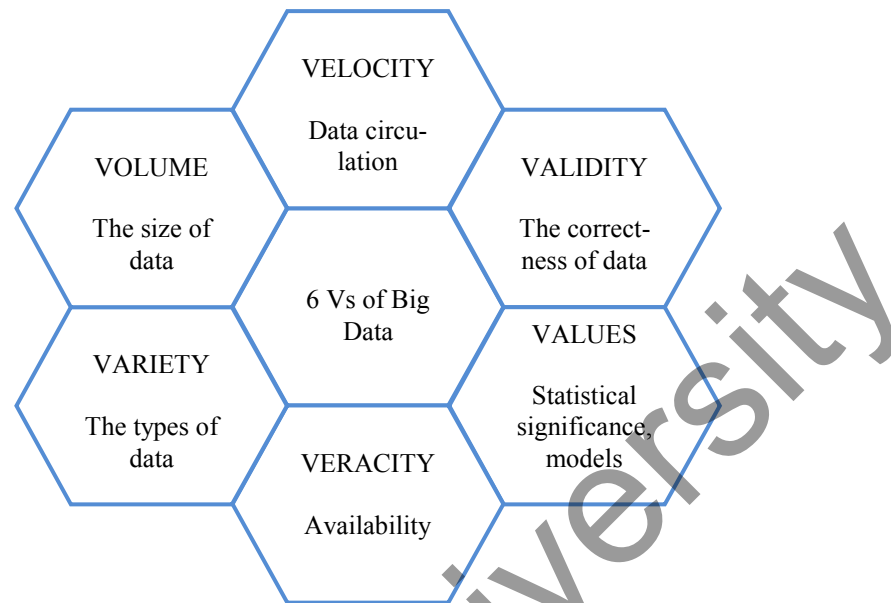


Figure 4. 6 Vs of Big Data

Note – Compiled by authors on the basis of (Daniel, 2017)

Big Data Analytics is employed at least in two directions such as a new course at universities, and a tool in public administration of education. In the case of many studies, Big Data Analytics helps to retain students at higher education institutions. The problem of dropping courses occurred at universities around the world. In the case of Kazakhstan, the drop of courses at universities needs to be studied as a gap for further research.

In summary, Big Data is evolving and its concept of theoretical framework is building up too. Particularly, Big Data and its analytics study in education in Kazakhstan are relevant. The main reason of that is the lack of research outputs demanded by market. It can be also explained by a lag of real data use in education.

### Conclusions

Nowadays, automated activities are everywhere in the world exemplifying the New Trends ranging from camera or scan usage, satellite TVs, distance education to the use of robots.

To comprehend the learning patterns and behaviors of students, to make strategic policies, the special institutions or governments are implementing EDM/LA. Educational Data Mining (EDM) and Learning Analytics (LA) described some similar tools, which are beneficial and useful to researchers, professors, administrators, and students by investigating the behavioral patterns of the students (Ray & Saeed, 2018).

More specifically Annapoorna Manohar et al. (2019) focused on assuming the university ranks employing forecasting about mining technique. For instance, as a model of Student Performance – the prediction and analysis showed several factors which pre-empted the probability of a success of a university student. Microsoft SQL Server Data Mining Add-ins Excel 2008 is a tool which can give assumptions on which universities can have special corrective measures to improve the quality of education and increase the benefits of faculty towards people. MongoDB; Hadoop; MapReduce; Orange; Weka; SAP HANA are top open source tools in Data science (Sin & Muthu, 2015).

These technologies mentioned above are useful to apply for big data. There are many big data sources in global and local contexts. These databases open new challenges and opportunities for development.

In summary, Big Data technologies and sources are used in all spheres of the economy including education sector, where the information is becoming large.

The theoretical framework of Big Data and its analytics continues to evolve. There is a lack of criticisms of theory and models limited with startups. Exact theories and models do not exist yet. In particular, Big Data Analytics in education is just absorbing new technologies to use.

COVID-19-related global crisis enhanced the use of Big Data Analytics in terms of online teaching and working at home. Education policy makers over the world have been forced to use different big data technologies towards data-driven decisions. The increase of data size, data driven decisions, and development of new technologies lead to over popularity of Big Data Analytics for all industries including education system.

### References

- Ashraf, F. (2020). Rise of Big Data Analytics and the Role of Human Resource Management. *American Journal of management*, 20 (2), 90–95. <https://doi.org/10.33423/ajm.v20i2.3001>
- Berendt, B., Littlejohn, A., Kern, P., Mitros, P., Shacklock, X., & Blakemore, M. (2017). Big data for monitoring educational systems. *Publications Office of the European Union*, Luxembourg, 20–25. <https://doi.org/10.2766/38557>
- Chad, L., Springer, J., Seliger, C., & Na, L. (2017). Impacting Big Data analytics in higher education through Six Sigma techniques. *International Journal of Productivity and Performance Management*, Vol. 66 No. 5, 15–21. <https://doi.org/10.1108/IJPPM-09-2016-0194>
- Cprime An Alten Company (n.d.). When Is Big Data Big? *cprime.com*. <https://www.cprime.com/resources/blog/when-big-data-big/>
- Daniel, B.K. (2017). Big Data in Higher Education: The Big Picture. *Big Data in Higher Education: The Big Picture. Big Data and Learning Analytics in Higher Education*, 19–28. DOI: 10.1007/978-3-319-06520-5\_3
- Drigas, A., & Leliopoulos, P. (2014). The Use of Big Data in Education. *IJCSI International Journal of Computer Science Issues*, Vol. 11, Issue 5, No. 1, pp. 58–63. [https://www.researchgate.net/publication/274890131\\_The\\_Use\\_of\\_Big\\_Data\\_in\\_Education](https://www.researchgate.net/publication/274890131_The_Use_of_Big_Data_in_Education)
- Lane, J. (2019). Big Data: The role of education and training. *Journal of Policy Analysis and Management*, 723–750. <https://doi.org/10.1002/pam.21922>
- Manohar, A., Gupta, P., Priyanka, V.M., & Uddin, M.F. (2016). Utilizing Big Data Analytics to Improve Education University of Bridgeport. Bridgeport, CT USA. <https://scholarworks.bridgeport.edu/xmlui/bitstream/handle/123456789/1616/Pooja%20Gupta%20-%20ASEE.pdf?sequence=1&isAllowed=y>
- Mergel, I. (2016). Big Data in Public Affairs Education. *Journal of Public Affairs Education*, 2, 231–248. <https://doi.org/10.1080/15236803.2016.12002243>
- McAfee, A., & Brynjolfsson, E. (2012). Big Data's Management. *Harvard Business Review*, 90 (10), 60–6, 10–15. <https://hbr.org/2012/10/big-data-the-management-revolution>
- McNeely, C.L., & Hahn, Jo. (2014). The Big (Data) Bang: Policy, Prospects, and Challenges. *Review of Policy Research*, Vol. 31, 304–310. <https://doi.org/10.1111/ropr.12082>
- Miftachul, H., Anshari, M., Shahrill, M., & Almunawar, M. (2016). Innovative eaching In Higher Education: The Big Data Approach. *TOJET: The Turkish Online Journal of Educational Technology, Special Issue for INTE*, 1210–1216. <https://doi.org/10.13140/RG.2.1.1267.6087>
- Mohsen, A., John, S., & Derek, S. (2018). Opportunities and Challenges for Big Data Analytics in American Higher Education-A Conceptual Model for Implementation. *Industry and Higher Education*, 32(1), 2–12. <https://doi.org/10.13140/RG.2.1.1267.6087>
- Ray, S., & Saeed, M. (2018). Applications of Educational Data Mining and Learning Analytics Tools in Handling Big Data in Higher Education, Researchgate, 2–10. [https://doi.org/10.1007/978-3-319-76472-6\\_7](https://doi.org/10.1007/978-3-319-76472-6_7)
- Sekli, M., & Vega, I. (2021). Adoption of Big Data Analytics and its impact on Organizational performance in Higher Education Mediated by Knowledge Management. *Journal of Open Innovation: Technology, Market, and Complexity*, 7(4), 221–231. <https://doi.org/10.3390/joitmc7040221>
- Shabbir, M., & Gardezi, S. (2020). Application of big data analytics and organizational performance: The mediating role of knowledge management practices. *Journal of Big Data*, 7, 47–57. <https://doi.org/10.1186/s40537-020-00317-6>
- Sin, K.W., & Muthu, L. (2015). Application of Big Data in Education Data Mining and Learning Analytics. *ICTACT Journal on soft computing: special issue on soft computing models for Big Data*, 05, Issue: 04 Analytics, 1035–1049. <https://doi.org/10.21917/ijsc.2015.0145>
- Tulasi, B. (2013). Significance of Big Data and Analytics in Higher Education. *International Journal of Computer Applications (0975 – 8887)*, 68, No.14, 21–23. <https://doi.org/10.5120/11648-7142>
- Weiss, S., & Indurkha, N. (1998). Predictive Data Mining: A Practical guide, Morgan Kauffman. *Elsevier Science & Technology Books*, 228.
- Zachary, A. (2017). Big data in education and the models that love them. *Current Opinion in Behavioral Sciences*. Vol. 18, 113–117. <https://doi.org/10.1016/j.cobeha.2017.11.006>

Д.К. Ильясов, Г.А. Райханова

## Білім беруді мемлекеттік басқаруда Big Data технологияларын пайдаланудың қиындықтары

### Аңдатпа

**Мақсаты:** Бүгінгі күні үлкен деректердің таралуы деректерді жинау мен талдаудың жаңа мәселелерін қарастыру арқылы жылдамырақ қозғалуды талап етеді. Зерттеудің негізгі мақсаты — білім беруді мемлекеттік басқаруда Big Data технологияларын қолданудағы негізгі мәселелерді анықтау болып табылады.

**Әдісі:** МШТҚ (мүмкіндіктер, шектеулер, трендтер және қиындықтар) Big Data Analytics теориялық даму бағытын қамтамасыз етеді. Сонымен қатар, бұл мақалада жетекші университеттерде үлкен деректерді талдауды халықаралық қолдану негізінде білім беруде деректерді талдауды қолданудың тұжырымдамалық негізі ұсынылған. Big Data Analytics білім беру жүйесіндегі мемлекеттік басқаруда және білім беру мекемелеріндегі оқу курсы ретінде екі рөл атқарады. Бүгінгі таңда деректердің көлемі аналитиканың артқа қалуымен айтарлықтай өсуде. Бұл зерттеу үлкен деректерді талдау мен білімді басқару арқылы ұйымның жетістігі арасындағы байланысты бекітті. Осы гипотезаларды тексеру үшін Нархоз университетінің 65 қызметкерінен сауалнама алынды. Құрылымдық модельдеу және растайтын факторлық талдау білімді басқарумен делдал болған ұйымдастырушылық тиімділік үшін үлкен деректерді талдауды бейімдеудің маңыздылығын тексеру үшін қолданылады.

**Қорытынды:** Big Data Analytics теориялық тұжырымдамалары әлі де жаңа мәселелер мен сын-тегеуріндерге қарсы заманауи технологияларды енгізу арқылы қалыптасуда. Атап айтқанда, білім берудегі Big Data Analytics қолданудың одан әрі даму мәселелері бар, олар да шешімдер қабылдауға және мемлекеттік басқару міндеттерін шешуге жұмсалады.

**Тұжырымдама:** Нәтижесінде осы зерттеуде Қазақстанның жоғары оқу орындарында білім беруді басқаруда Big Data Analytics қолданудың басым бағыттары белгіленген.

**Кілт сөздер:** Big Data, Big Data технологиялары, болжауды талдау, студент үлгерімінің менеджменті.

Д.К. Ильясов, Г.А. Райханова

## Проблемы использования технологий больших данных в государственном управлении образованием

### Аннотация

**Цель:** Основной целью данного исследования является выявление ключевых проблем в использовании технологий больших данных в государственном управлении образованием.

**Методы:** OLTC (возможности, ограничения, тенденции и проблемы) обеспечивают путь теоретического развития аналитики больших данных. Кроме того, в этой статье предложена концептуальная основа применения аналитики данных в образовании на основе международного использования аналитики больших данных в ведущих университетах. Аналитика больших данных играет двоякую роль в государственном управлении в системе образования и как учебный курс в образовательных учреждениях. Сегодня объем данных стремительно растет, высокими темпами опережая реальность за счет отставания их аналитики. В данном исследовании утверждена взаимосвязь между аналитикой больших данных и успехом организации посредством менеджмента знаний. Для проверки этих гипотез были опрошены 65 сотрудников университета Нархоз. Структурное моделирование и подтверждающий факторный анализ применяются для проверки значимости адаптации аналитики больших данных для организационной эффективности, опосредованной управлением знаниями.

**Результаты:** Теоретические концепции аналитики больших данных все еще формируются путем внедрения современных технологий с последующими новыми проблемами и вызовами. В частности, применение аналитики больших данных в образовании имеет свои вопросы дальнейшего развития, затрагиваемые также для принятия решений и решения задач государственного управления.

**Выводы:** В результате в данном исследовании обозначены приоритетные направления применения аналитики больших данных в управлении образованием в высших учебных заведениях (вузах) Казахстана.

**Ключевые слова:** большие данные, аналитика больших данных, аналитика прогнозирования, менеджмент успеваемости студента.