Model of Intelligent Massive Open Online Course Development

Gulmira Bekmanova[1,2(✉)], Assel Omarbekova[1], Zulfiya Kaderkeyeva[1], and Altynbek Sharipbay[1,2]

[1] L.N. Gumilyov Eurasian National University, Nur-Sultan, Kazakhstan
gulmira-r@yandex.kz, omarbekova@mail.ru, zulfiya_83@mail.ru, sharalt@mail.ru
[2] Nuclear University MEPhI, Moscow, Russia

Abstract. The relevance of development and extension of massive open online courses (MOOC) got a new wave of development due to the coronavirus pandemic. The importance and necessity of MOOCs will be increasing, however, intelligent systems changed qualitatively since the development of first MOOCs. The intelligent MOOC development with using Kazakh language thesaurus approach is suggested in this paper. The model of intelligent MOOC suggests laying its intellectuality at its designing, using the knowledge base, ontological model of discipline, and their relevant question-answer system and intelligent search. The separate important part of each such MOOC is the intelligent assessment of knowledge and achievement of training's announced results. The suggested MOOC model makes it more effective means for distance, blended and any e-learning. The intelligent MOOC possesses a possibility of its using in e-learning systems without a tutor.

Keywords: Intelligent massive open online course · Intelligent assessment of knowledge · Control over learning results achievement · Ontological model · Kazakh language thesaurus assessment of human language plaintexts

1 Introduction

Since their establishment, the massive open online courses (MOOCs) were meant to revolutionize and democratize higher education [1, 2]. The American publisher, the New York Times called the year of 2012 as a year of massive open online courses. The number of users is consistently and steadily growing since that time. As per 2017 results, the number of online courses participants was over 81 million people. The total number of MOOCs certificates issued for successful learning is also steadily growing [3].

The most popular services are the US platforms Coursera, edX and Udacity, as well as British FutureLearn. Many countries created their national online platforms: France Université Numérique (FUN) – in France, MiriadaX – in the Latin American countries, EduOpen – in Italy, XuetangX – in China, SWAYAM – in India, National platform for open education in Russia (NPOE), Educational platform BilimLand in Kazakhstan.

In 2020, the pandemic and lockdown regime caused a fast growth in number of online courses users. According to the GetCourse platform, hosting distance courses for

© Springer Nature Switzerland AG 2020
O. Gervasi et al. (Eds.): ICCSA 2020, LNCS 12250, pp. 271–281, 2020.
https://doi.org/10.1007/978-3-030-58802-1_20

hundreds of online schools, the sales of courses already increased by 20% in the second half of March compared to the first half of the month. The number of users tried to organize their own courses herein increased by 30% [4].

Over 1 million people completed online additional medical training courses to help coronavirus infected patients [5].

During the quarantine, many online training services, such as Coursera, Khan Academy, GeekBrains, Arzamas made great number of courses free, which also leads to increase of MOOCs users [6].

At the moment the online course should not only provide high-quality content, but comply with current trends and new technologies as well.

Here we list the main trends of online learning which are relevant in 2020:

1) Individual learning path. The modular training courses are becoming increasingly popular where the participant may choose personal learning path, which meets his needs.
2) Mobile training. It is important for participants to have an access not only with a PC, but via mobile device as well.
3) Micro training. With the modern man's limited free time, the educational videos tend to shorten by their duration.
4) Online training with extensive feedback. In order to learn effectively and quickly achieve results, users are willing to pay extra for the feedback function.
5) Community trend and networking. More and more users attend courses not only to obtain knowledge, but also to get acquainted to their like-minded people.
6) Practice-oriented training. Users are more interested in development of practical skills than in obtaining theoretical knowledge.
7) Traineeships and employment as part of an online course. Those educational products that provide traineeships in large companies for students or help the most successful students in finding work by a new, just obtained profession, win the competitive struggle.
8) Blended training. The courses which combine offline and online formats are in demand.
9) Multi-format training. Online courses, where information is presented in various formats, are in demand. Under this approach, information is digested in better way, and the course becomes more replete.
10) Gamification. Perhaps it is the most controversial trend among the mentioned. On the one hand, game mechanics in training work perfectly and help to maintain interest in classes and complete online courses. But - only in case if game mechanics are relevant to educational tasks and help to achieve learning goals [7].

The MOOC development and extension received a new strong impulse for development due to the coronavirus pandemic, and the importance and necessity of MOOCs will be only growing predictably. But the MOOCs existing on the market most often do not use artificial intelligence methods at development of MOOCs. The paper proposes an approach for development of intelligent MOOC from its designing to use artificial intelligence in assessment of student's knowledge that will make it more effective means of distance, blended and any e-learning. These are exactly the

modern requirements for MOOCs. There are many various models of MOOCs. The considered variant of MOOC contains the following didactic materials:

- Glossary
- Presentation
- Webinar
- Lecture materials
- Quiz questions
- Practical and/or laboratory tasks
- Individual work tasks
- Final exam materials.

For the MOOCs development, all materials are presented as an ontological model, which is an effective tool for representing knowledge [links]; fuzzy methods for assessing students' knowledge are used to evaluate their knowledge level, which are disclosed in details in the papers [links].

Labor costs for the development of single MOOC are at least 100 h and include the following main stages:

1. Formulation of training results
2. Elaboration of structure and content of MOOC
3. Development of assessment system
4. Creation of MOOC video-content
5. Allocation of MOOC materials on a platform
6. MOOC expertise

Stages of intelligent massive open online course development remain but the content of stages changes.

The formulation of training results is carried out with consideration of the training form, thus the achievement of training results may differ between distant training, blended and traditional one (within the traditional training, MOOC is used as an auxiliary tool, and as the main one in distant training, in blended training, there are high-quality practical and/or laboratory exercises that take place in face to face manner as a rule). The most thoroughly and qualitatively developed assignments, questions, feedback and assessment, which verify the achievement of training results are needed for their guaranteed achievement in distance learning.

2 Related Works

There are articles studying the use of video-supported technology to facilitate learning's latest trend in the e-learning [8], perspectives of e-learning development [9], in which it is pointed out that despite the difference between educational traditions the resulting analysis brings to the surface not only how the values that underpin e-learning development in each region differ but also how specific perspectives influence the respective fields. The researchers acknowledge these differences but also remark on the historic and contemporary symbiosis that has endured even in this relatively new field of e-learning.

E-learning systems actively use artificial intelligence models and methods [10–13]. In the paper [14], authors develop a model to automatically label a post based on the first phase of the interaction analysis model. This allows instructors to automatically identify whether students are stating opinions, clarifying details, or engaging in activities such as providing examples to peers. Manuscript [15] presents a novel approach for developing an affective tutoring system for the MOOCs, which is called ATS-MOOCs. Such system can easily help students to improve their learning performance by recognizing their affective states and then adapting the MOOC content accordingly. Paper [16] shows multidimensional deep learner model combining these features with natural language processing. To illustrate presented method, authors used a benchmark dataset of 29598 posts, from three different academic subject areas. The findings highlight that the combined, multi-dimensional features model is more effective than the text-only (NLP) analysis, showing that future models need to be optimised based on all these dimensions, when classifying urgent posts.

The ontology is successfully applied in modeling systems for various types of knowledge. For instance, ontological models for knowledge presentation are suggested in the paper [17]. Besides, the presentation of ontology models allows converting of a model into an RDF scheme, which will be used for generation of reference reviews [18]. It is disclosed in the paper [19] how semantic hypergraphs are used to create ontological models of morphological rules in the Kazakh language. The ontology of professional competencies and e-CF are considered in the paper [20], it proposes [21] the competence correspondence for training programs to national and international requirements. The modular competencies, ontological approaches and hierarchy analysis are used in the work to compare the competencies of the training program, professional standards and e-CF. There has been built an ontology which comprises competence data and determines their relation.

The elaboration of systems able to make logical conclusions from the knowledge base is rapidly developing all over the world. The extraction of existing and generation of new knowledge and statements from the existing knowledge base is carried out with specialized engines intended to reason and make logical conclusions. In the paper [22], the authors propose to implement RDF to represent knowledge and use a data-query language to extract knowledge. In [23], developers describe methods for a knowledge base processing, extracting triplets, and generating queries.

In the articles [24, 25], we described the developed system for the automatic generation of test questions based on the ontological model. The development of test questions for students' knowledge assessment requires no-nonsense approach. The test tasks of high quality are developed thoroughly in a way that they would not be ambiguous or unclear for the tested students. Various methodologies were investigated under test tasks development as well as deeply studied Bloom's taxonomy. In accordance with taxonomy, the students' skills are divided into six categories, knowledge, comprehension, application, analysis, synthesis and assessment. The most primitive of the skills are knowledge and comprehension, and the most advanced are analysis, synthesis and assessment. It has been decided to generate test tasks to check low level skills of knowledge and comprehension in the process of development in a way that a system able to make full argument would be needed for generating questions for analysis and assessment. Special verbs were determined in accordance with Bloom's

taxonomy to develop test tasks. Some of the verbs are for knowledge and definition, to give definition, to enumerate, to name, to explain, etc. Low level skills may be estimated with the mentioned verbs while verbs for classification, finding alternative and description of advantages are used to estimate students' skills of higher level. The developed system generates test questions to assess the skills of low-level students.

The system, according to the taxonomy of B.S. Bloom, automatically generates assignments for assessing knowledge and comprehension of educational material. The development of knowledge bases for intelligent systems is shown in paper [26].

3 Formulation of Training Results

The formulation of training results is carried out with consideration of the training form, thus the achievement of training results may differ between distant training, blended and traditional one (within the traditional training, MOOC is used as an auxiliary tool, and as the main one in distant training, in blended training, there are high-quality practical and/or laboratory exercises that take place in face to face manner as a rule). The most thoroughly and qualitatively developed assignments, questions, feedback and assessment, which verify the achievement of training results are needed for their guaranteed achievement in distance learning.

4 Discipline's Ontological Model

The ontological model of discipline is developed to create the MOOC's structure and content. The ontological model is built in Protégé tool [27]. It was developed at Stanford University in collaboration with the University of Manchester.

The ontology is a powerful and widely used modeling tool for relations between objects belonging to different fields of a subject. The ontology defines a set of representative primitives in the context of computer and information sciences with which we can model a field of knowledge or a discourse. The representative primitives are usually classes (or sets), attributes (or properties), and relationships (or relationships between members within a class). The definitions for representative primitives include information about their meaning and limitations of their logically concerted application.

This formalism determines the "O" ontology as triple (V, R, K), where V – is a set classes for the subject field, R – is a set of relationships between the classes, and K – is a set of attributes within the field [17, 28].

The ontological model consists of the discipline topics and the glossary, each of the discipline topics includes glossary's basic concepts, control questions are developed on the topic linked to the glossary, as well as the assignments to verify the achievement of results by each of the topics, a knowledge base of questions and answers by the ontology is created in human language, the system can self-learn within the framework of a set discipline (see Fig. 1).

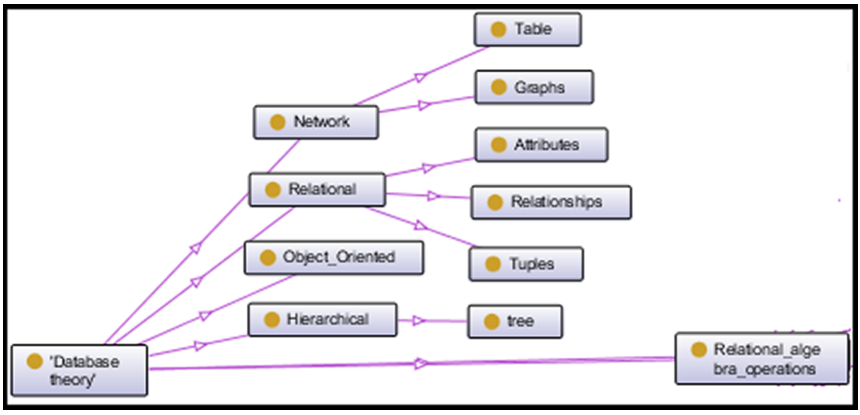

Fig. 1. "Databases theory" discipline's ontological model

The figure (see Fig. 2) shows an ontology fragment by topic of relational algebra operation, "Databases theory" discipline, linked to discipline's competences.

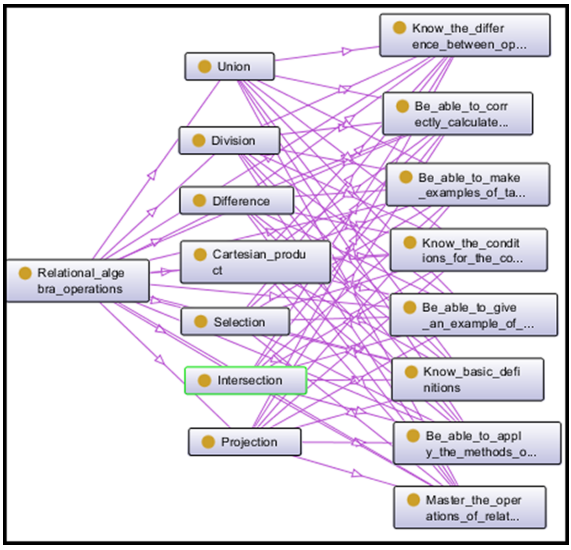

Fig. 2. A fragment of "Databases theory" discipline's one topic ontological model

Such concept allows:

1) To link fragments of video-lecture with glossary's key-words and to rewind the lecture by the key-words.
2) To find answers to the lecture questions automatically in the video-lecture.
3) To carry out the intelligent search in a MOOC within the discipline in the human language [29].

5 Development of the Assessment System

It is suggested to use assessment system based on the open tests for a certain discipline. A student writes his answer which is then compared with a reference etalon answer based on a fuzzy binary relation, set between a student's answer and reference ones synthesized from the knowledge base of a set discipline in accordance with test questions. Then it becomes possible to figure out the competence marks by all modules on assumption of the marks obtained by the disciplines within the modules. As a result we may assess the achievement of training results by the whole educational program, which is calculated as an average from the compulsory results achievement according to the disciplines of the educational program.

Let P – be a finite and nonempty set representing the knowledge base of a given discipline; Q – a finite and nonempty set representing a base of test questions in a given discipline; T – a finite and non-empty set representing a base of standard answers to test questions in a given discipline;

$f: Q \to T, f(q) = t, q \in Q, t \in T$ (4) – a function to generate answers to asked questions; S – a finite set representing the knowledge base of student answers to test questions in the given discipline.

'Mark' $\rightleftharpoons$

where $\mu_{S \subseteq T}: S \times T \to]0, 1[$ (12) fuzzy relation function.

In order to calculate the elements of fuzzy relation $S \subseteq T$ it is needed to set members S и T, which are the range of class individuals, range of class properties and the range of semantic arcs incidental to the class.

Intelligent system for evaluating the level of formation of professional competencies of students is described in papers [30–32] in details.

6 Creation of the MOOC Video-Content

Types of video-lectures and video-lessons:

1. Lecturer's video ("talking head"). This is the least productive and didactically ineffective form of distance training.
2. Live video ("for skippers"). It is a video of university lectures recorded directly indoors. For example, the Massachusetts Technology Institute arranged one of the biggest video-lecture collections recorded in lecture hall.
3. Studio video-lectures. Such video records are well edited and directed. All flakes, teacher's "goofs" are deleted. Very often such video-lesson is accompanied with demonstration of images, video clips and very close to a documental film by its level. Currently we are recording video-lectures of math disciplines in such format.
4. Slide films. The footage in this case is in a key place and accompanied by an off-screen comments made by the teacher or speaker. Such film is as close as possible to a documental educational film. Such type of video-lectures is ineffective and may be used as teacher's auxiliary material.

5. Interactive video-lectures and video-lessons. The teacher's monologue is accompanied by slides, video fragments, assignments. It is the most preferable variant of video-lectures (see Fig. 3).

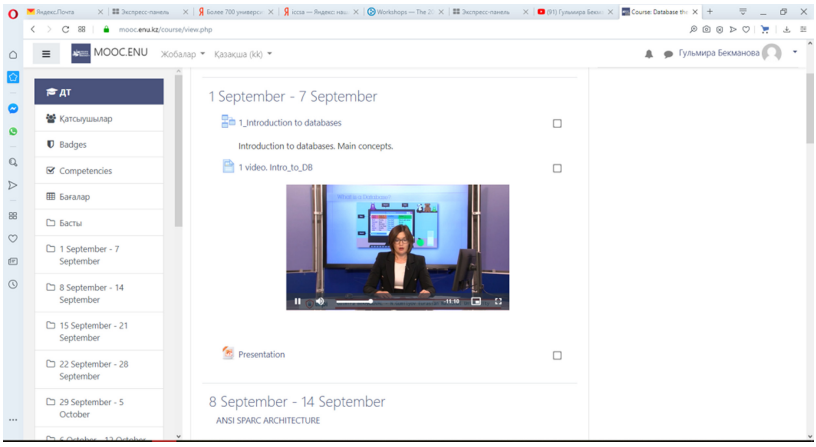

Fig. 3. Database theory MOOC fragment

The figure shows MOOC fragment for Database theory discipline, at the moment the work is being done to implement the open tests based assessment.

7 Allocation of MOOC Materials on a Platform

This stage does not have any difference by MOOC allocation technology. The requirements for a platform itself are changing. It is the platform that carries out the implementation of all functions of intelligent MOOC. Such platform satisfies the following requirements:

1) verification of results achievement by discipline through the connection between content and training results, assignments, open tests based on thesaurus.
2) mechanisms of intelligent assessment of students' knowledge and competences.
3) Ensuring of feedback with students.

8 MOOC Expertise

The expertise includes the standard MOOC expertise for the content and didactic materials quality, technical expertise of video-lectures, and pedagogical expertise of MOOCs. The verification of the MOOC intelligent function operation correctness is engaged additionally for: intelligent search, system's verification of open tests, verification of questions and assignments regarding setting up of competencies and training results.

9 Conclusion

In this work, the model of intelligent massive open online course development has been proposed. The discipline content is presented in a shape of ontological model, which provides possibility to create connections between discipline topics, glossary, assignments and questions. Thus, the intelligent search, questions and answers in human language during learning, verification of achievement of training results through assignments, as well as students' open answers of students become possible. Such approach, despite of being more time consuming at creation of intelligent MOOC, will allow to educate a student with minimal involvement of a teacher in the training process after setting the system. The system has development perspectives as an educational system without teacher's involvement. At the same time, with a huge knowledge base it may be successfully used to guarantee the verification of the training results achievement.

Acknowledgments. The work was supported by the grant financing for scientific and technical programs and projects by the Ministry of Science and Education of the Republic of Kazakhstan (Grant No. AP05132249, 2018–2020).

References

1. Belanger, Y., Thornton, J.: Bioelectricity: A quantitative approach Duke University's first MOOC (2013)
2. Haggard, S.: The maturing of the MOOC. Department for Business Innovation & Skills (2013)
3. Chuang, I., Ho, A.: HarvardX and MITx: Four Years of Open Online Courses – Fall 2012-Summer 2016 (2016)
4. https://www.vedomosti.ru/management/articles/2020/04/08/827543-onlainovie-kursi
5. https://xn–80aesfpebagmfblc0a.xn–p1ai/news/20200413-1847.html
6. https://1prof.by/news/stil-zhizni/attrakcion-shhedrosti-iz-za-koronavirusa-mnogie-servisy-stali-besplatnymi/
7. https://edmarket.ru/blog/top-trends-of-e-learning
8. Chen, Y.: A study of implementing and evaluating an interactive video on demand. New Educ. Rev. **28**(2), 115–125 (2012)
9. Hillen, S., Landis, M.: Two perspectives on e-learning design: a synopsis of a U.S. and a European analysis. Int. Rev. Res. Open Dist. Learn. **15**(4), 199–225 (2014)
10. Sharipbay, A., Barlybayev, A., Sabyrov, T.: Measure the usability of graphical user interface. New Advances in Information Systems and Technologies. AISC, vol. 444, pp. 1037–1045. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-31232-3_98
11. Sharipbay, A., Omarbekova, A., Seifullina, A., Zakirova, A.: Creating intelligent electronic textbooks based on knowledge management. In: The World Congress on Engineering 2014, vol. 1, pp. 224–227 (2014)
12. Sharipbay, A., Omarbekova, A., Nurgazinova, G.: Creating intelligent electronic textbooks in Kazakhstan. In: 6th International Conference on Education and New Learning Technologies. Barcelona/Spain, pp. 2926–2933 (2014)

13. Nurgazinova, G., Golenkov, V., Sharipbay, A., Omarbekova, A., Barlybayev, A.: Designing of intelligent reference system for algebra based on the knowledge database. In: International Conference on Control, Automation and Artificial Intelligence, pp. 230–235 (2015)
14. Pillutla, V.S., Tawfik, A.A., Giabbanelli, P.J.: Detecting the depth and progression of learning in massive open online courses by mining discussion data. Technol. Knowl. Learn. (2020). https://doi.org/10.1007/s10758-020-09434-w
15. Soltani, M., Zarzour, H., Babahenini, M.C., Chemam, C.: An affective tutoring system for massive open online courses. In: Bouhlel, M.S., Rovetta, S. (eds.) SETIT 2018. SIST, vol. 146, pp. 202–211. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-21005-2_20
16. Alrajhi, L., Alharbi, K., Cristea, A.I.: A multidimensional deep learner model of urgent instructor intervention need in MOOC forum posts (2020). https://doi.org/10.1007/978-3-030-49663-0_27
17. Yergesh, B., Mukanova, A., Sharipbay, A., Bekmanova, G., Razakhova, B.: Semantic hyper-graph based representation of nouns in the Kazakh language. Computacion y Sistemas **18**(3), 627–635 (2014)
18. Bekmanova, G., Sharipbay, A., Omarbekova, A., Yelibayeva, G., Yergesh, B.: Adequate assessment of the customers actual reviews through comparing them with the fake ones. In: Proceedings of 2017 International Conference on Engineering and Technology, pp. 1–4 (2018)
19. Mukanova, A., Yergesh, B., Bekmanova, G., Razakhova, B., Sharipbay, A.: Formal models of nouns in the Kazakh language. Leonardo Electron. J. Pract. Technol. **13**(25), 264–273 (2014)
20. Bazarova, M., Zhomartkyzy, G.: Ontological model of professional competences for ICT-specilaists 6D070300. Vestnik KazNITU. №1 (119), pp. 321–328 (2017)
21. Bazarova, M., Zhomartkyzy, G., Wojcik, W., Krak, Yu.: Construction of individual trajectories of training specialists in the field of information and communication technologies. J. Autom. Inf. Sci. **49**(10), 36–46 (2017)
22. Liu, BL., Hu, B., Wang, XY., Zheng, L.: A RDF-based knowledgebase implementation for E-learning. New Horizon in Web-based Learning. In: 3rd International Conference on Web-Based Learning, pp. 45–50 (2004)
23. Wang, L., Zhang, Yu., Liu, T.: A deep learning approach for question answering over knowledge base. In: Lin, C.-Y., Xue, N., Zhao, D., Huang, X., Feng, Y. (eds.) ICCPOL/NLPCC 2016. LNCS (LNAI), vol. 10102, pp. 885–892. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-50496-4_82
24. Omarbekova, A., Sharipbay, A., Barlybaev, B.: Generation of test questions from RDF files using PYTHON and SPARQL. J. Phys. Conf. Ser. **806**(1), 012009 (2017)
25. Omarbekova, A., Nurgazinova, G., Sharipbay, A., Barlybayev, A., Bekmanova, G.: Automatic formation of questions and answers on the basis of the knowledge base. Paper presented at the Proceedings of 2017 International Conference on Engineering and Technology (2018)
26. Yelibayeva, G., Mukanova, A., Sharipbay, A., Zulkhazhav, A., Yergesh, B., Bekmanova, G.: Metalanguage and knowledgebase for Kazakh morphology. In: Misra, S., et al. (eds.) ICCSA 2019. LNCS, vol. 11619, pp. 693–706. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-24289-3_51
27. https://protege.stanford.edu/
28. Yergesh, B., Bekmanova, G., Sharipbay, A., Yergesh, M.: Ontology-based sentiment analysis of kazakh sentences. In: Gervasi, O., et al. (eds.) ICCSA 2017. LNCS, vol. 10406, pp. 669–677. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-62398-6_47

29. Sharipbayev, A., Bekmanova, G., Yergesh, B., Mukanova, A., Buribayeva, A.: Semantic retrieval of information in the Kazakh language in elibraries. J. Int. Sci. Pub. Educ. Alter. **10**, part 1, 108–115. (2014)
30. Zulfiya, K., Gulmira, B., Altynbek, S.: Ontological model for student's knowledge assessment. Paper presented at the ACM International Conference Proceeding Series (2019)
31. Zulfiya, K., Gulmira, B., Altynbek, S., Assel, O.: A model and a method for assessing students' competencies in e-learning system. Paper presented at the ACM International Conference Proceeding Series (2019)
32. Barlybayev, A., Kaderkeyeva, Z., Bekmanova, G., Sharipbay, A., Omarbekova, A., Altynbek, S.: Intelligent system for evaluating the level of formation of professional competencies of students (2020)