JSC "INTERNATIONAL INFORMATION TECHNOLOGY UNIVERSITY»

UDC: 004.89

SMAIYL ASSEL MARALBAIKYZY

Models and methods of intellectual information and training system

Thesis

Presented for the Degree of Doctor of Science in Information System
(Degree code: 6D070300) Department of Information System

Supervisor:
Doctor of Technical Sciences,
Professor Kuandykov A.A.

External supervisor:
PhD, Marco Brambilla

Republic of Kazakhstan
Almaty, 2021

CONTENT

REGULATORY REFERENCES

This dissertation uses references to the following standards:

- «Instructions for the preparation of a dissertation and an abstract", Higher Attestation Commission of the Ministry of Education and Science of the Republic of Kazakhstan dated September 28, 2004 No. 377-3 y.
- GOST 7.32-2001-Report on research work. Structure and design rules.
- GOST 7.1-2003. Bibliographic record. Bibliographic description. General requirements and rules of compilation.
- ST RK 34.014 -2002-Information technology. A set of standards for automated systems. Automated systems. Terms and definitions.
- ST RK 34.005-2002 Information Technology. Basic terms and definitions-introduced for the first time.
- ST RK. 34.015-2002 Information Technology. Set of standards for automated systems. Terms of reference for creating an IP-introduced for the first time.
- ST RK 34.027-2006 Information Technologies. Classification of software tools-introduced for the first time.
- ST RK 34.014 -2002 Information Technology. A set of standards for automated systems. Automated systems. Terms and definitions.

DEFINITIONS

In this dissertation work, the following definitions are used:

Adaptive system is a system that automatically changes the data of its functioning algorithm and (sometimes) its structure in order to maintain or achieve an optimal state when external conditions change.

An intelligent system is a technical or software system capable of solving problems that are traditionally considered creative, belonging to a specific subject area, knowledge about which is stored in the memory of such a system.

An ontology is a system of concepts of a certain subject area, which is represented as a set of entities connected by various relationships. Ontologies are used to formally define the concepts and relationships that characterize a particular area of knowledge.

Distance learning, also called distance education, is the education of students who may not always be physically present at organization. Traditionally, this usually involved correspondence courses wherein the student corresponded with the organisation via mail, or other online tools.

Information management is processes related to the organization of creation, modification, destruction, search and storage of data.

Microlearning deals with relatively small learning units and short-term learning activities. The term is used in e-learning and related fields in the sense of learning processes in mediated environments.

Semantics is a branch of linguistics that studies the semantic meaning of language units. Semantic analysis is used as a learning tool.

Semantic proximity is a metric defined over a set of documents or terms, where the idea of distance between items is based on the likeness of their meaning or semantic content as opposed to lexicographical similarity.

SYMBOLS AND ABBREVIATIONS

IS – Information Systems

ICT – Information and Communications Technology

OWL - Web Ontology Language

RDF - Resource Description Framework

RDFs - Resource Description Framework Schema

SWOT - Strengths, Weaknesses, Opportunities and Threats

XML - Extenbile Markup Language

FES - Fact Extraction Scheme

INTRODUCTION

Relevance. Development of an educational resource that allows to organize an effective learning process, taking into account the semantic representation of resources within the Smart University.

The experience of many years of work of universities and centers of postgraduate education shows that in the system of continuing professional education, individual training is the main demand, and the main limitation in solving the problems of individual education is time. In the system of advanced training and retraining, the time factor is mainly manifested in the discrepancy between the terms of the specialist's need to study the material with the announced official schedule of classes in educational institutions. Another important aspect of this issue is the content of the announced training programs, which do not take into account the individual needs of future students.[1]

The learning process is the same for each student, and the material for each course is the same. For a strong learner, the content may be too simple and insufficient, and therefore ineffective, and for a weak learner, it may be complex and incomprehensible.

Thus, the search for alternative ways to individualize learning is still an urgent problem.[2]

Qualitatively new opportunities for self-training and improving professional knowledge are provided by new information technologies of distance learning (distance learning) using local and global networks, CD-ROMs, video cassettes, television cable and satellite broadcasting.

In contrast to the traditionally built courses of full-time and even more part-time training, the use of information technology opens the way for training directly in the workplace, which, if properly organized, allows you to individualize the process and allocate the necessary amount of time for training staff without any tangible interruptions in work.

The concept of computer-based learning is based on the principles of autonomy (self-management) of the process of cognition. Its implementation involves a new combination of training and control programs with a developed component of mutual moral responsibility of teachers and students. Autonomy in the educational process implies not only the independence of universities, but also the right of the student to choose an individual path of education within the framework of a multi-level education system. [3]

The object of the research is the application of information systems in the education sector.

The subject of research is methods and development of an educational resource.

Aim of research. System development using models and methods of an intellectual information system based on semantic data representation, taking into account the features of the trajectory and individualization of the learning process.

Main goals. To achieve the research goal, the following tasks were set and solved:

1) Review and analysis of existing information systems for educational resources;
2) Models and methods for building an educational resource;
   2.1) Formation of the content of the educational resource;
   2.2) Compilation of content based on semantic analysis;
   2.3) Determining the sequence of content;
   2.4) Management of the number of training elements in accordance with Microlearning;
3) Development of the system in the form of a web portal;
4) Software testing;
5) Software implementation at Astana IT University.

Scientific novelty of the study and the provisions submitted for defense

Models and methods of content formation, which are intended for any student, takes into account the characteristics of the user:

1) The methodology for constructing an educational resource is based on semantic analysis;
2) Establishing the coefficient of semantic proximity of a pair of topics based on the proposed hypothesis;
3) The location of expressions in the content (on slides, texts) is determined by the microlearning method;
4) The result of an experimental study of the developed system, showing an increase in the quality of education and efficiency due to the semantic representation of resources, also creates conditions for the digitalization of the educational process.

Theoretical and practical significance of the study

The theoretical and applied results obtained in this work can be used in the design and development of a system for information and training organizations. And also, the obtained scientific results can be developed as a theoretical basis and an applied basis for automating the learning process by taking into account the individual characteristics of each student (intellectualization).

The developed system with the use of models and methods of the intelligent information and training system is implemented in the university, "Astana IT University" LLP, which is currently actively used in training.

Publication of research results

The main results obtained during the dissertation work are published in 12 printed works, of which 4 articles are published in publications recommended by the Committee for Control in the Field of Education and Science of the Ministry of Education and Science of the Republic of Kazakhstan, 1 article is published in a publication indexed by the Scopus and Web of Science Core Collection databases, with an impact factor of 1.635, quartile for computer science and information system-Q3, percentile for computer science applications – 51, 4 articles are published in the collections of international foreign conferences (Canada, Spain, Ukraine, Germany),

6 articles are published in the collections of international scientific and practical conferences (Kazakhstan).

Approbation of research results and publications

The main provisions and scientific results of the work were reported and discussed at seminars of the Department "Information Systems" of the International University of Information Technologies and at international scientific and practical conferences: V All-Russian Congress of Young Scientists at the National Research University ITMO (Russia, St. Petersburg, 2016); International Scientific and Practical Conference "INNOVATIONS IN EDUCATION AND SCIENCE" dedicated to the 25th anniversary of Independence of the Republic of Kazakhstan and the 20th anniversary of the Suleiman Demirel University (Kazakhstan, Kaskelen, 2016); International scientific and practical Conference "The Path to Independence through Decades" dedicated to the 25th anniversary of Kazakhstan's Independence (Kazakhstan, Almaty, IITC, 2016); The 2nd International Conference "Information Technologies in Science and Industry 2016" (Kazakhstan, Almaty, IITC, 2016); International Scientific and Practical Conference "5th International Conference Science and Society – Methods and problems of practical application" (Canada, Vancouver, 2018); in the scientific and practical ONLINE conference "Digital Kazakhstan-the introduction of the mechanism of IT technologies in Technical and Vocational Education: features of the use of modular and credit learning technologies in the educational process" organized by the Republican Educational and Methodological Association (Kazakhstan, Almaty, 2018); "3rd International Conference on System Reliability and Safety (ICSRS 2018)" (Spain, Barcelona, 2018); 5th International Conference on "Digital technologies in science and industry-2019" (Kazakhstan, Almaty, 2019); 2021 IEEE Smart Information Systems and Technologies (SIST) (Kazakhstan, Nur-Sultan, 2021) and Dortmund International Research Conference 2021 (IRC 2021).

Structure and scope of work. The structure of the work is determined by the set goal and the sequence of solving the formulated tasks and is built on the problem-thematic principle. The dissertation consists of an introduction, four chapters, a conclusion and a list of references, including 77 names of the sources used.

1 ANALYSIS OF THE STATE OF PROBLEMS OF INTELLECTUAL INFORMATION AND TRAINING SYSTEMS

This section of the dissertation examines the importance and role of an intelligent information system for educational institutions. A review and analysis of modern approaches to the automation of educational processes is carried out, as well as a literature review. The features of the content of the discipline, as well as the sequence and number of educational elements are considered.

Information technologies occupy a central place in the process of intellectualization of society, the development of its educational system and culture. In addition, the use of training information tools has proved to be a very effective method for both self-education systems and systems of advanced training and retraining of personnel.

The processes of informatization of modern society and the closely related processes of informatization of all forms of educational activity are characterized by the processes of improvement and mass dissemination of modern information and communication technologies (ICT). Such technologies are actively used to transmit information and ensure interaction between the teacher and the student in modern systems of open and distance education. A modern teacher should not only have knowledge in the field of ICT, but also be a specialist in their application in their professional activities. The concept of technology includes the application of scientific and engineering knowledge to solve a practical problem. Then information technology can be considered the process of turning knowledge into an information resource. The purpose of information technology is the production of information for its subsequent analysis and decision-making on the basis of it to perform an action. Information and communication technologies (ICT) is a general concept that describes various devices, mechanisms, methods, and algorithms for processing information. With the help of network means of ICT, it becomes possible to have wide access to educational and methodological and scientific information, to organize operational consulting assistance, to model research activities, to conduct virtual training sessions (seminars, lectures) in real time.

Analyzing the social significance of information technologies, it seems appropriate to state the following:

- Modern information and telecommunications technologies make it possible to activate and effectively use the information resources of society, which are the most important strategic factor in its development.
- The development of civilization is taking place in the direction of the information society, in which the objects and results of the work of the majority of the employed population are no longer material values, but mainly information and scientific knowledge. At the same time, information technologies make it possible to optimize and, in many cases, automate the information processes taking place in society.

- Information processes are important elements of other more complex industrial or social processes. In this regard, information technologies are considered as components of the corresponding production or public technologies.

- Telecommunications technologies, being a part of information technologies, play an extremely important role in ensuring information interaction between people and organizations, as well as in the systems of preparation and dissemination of mass information [4].

World experience shows that the competitiveness of the national economy as a whole is associated with the development of information technologies. According to the World Economic Forum, the index of competitiveness of the national economy has a high level of correlation with the index of development in the countries of information and communication technologies. According to the 2012 World Economic Forum competitiveness rating of 142 countries, countries that actively develop information technologies are ahead of Kazakhstan in most of the indicators in terms of creating demand for information technologies - 51st place (USA-13, Germany - 19, India - 63, Egypt - 96), business conditions in information technologies - 104th place (USA - 21, Germany-38, India - 72).

One of the main reasons for the lag in IT development is the lack of innovative research and development centers and educational centers.

All of the above factors determine the need to initiate the development of Kazakhstan's own software development tools for the functioning of the telecommunications infrastructure and its components in the educational environment.. In addition, the process initiated in this direction will allow bringing the IT industry of the Republic of Kazakhstan to a qualitatively new level. [5]

1.1 Overview of existing information systems in the field of education

Before the pandemic in Kazakhstan, many universities did not have distance education. Only in some universities there was still online training only for the correspondence department or for the second higher education. Universities such as Narxoz, AlmaU, KazNU, University of International Business and others used distance education without adaptive learning. The course content is the same for users in accordance with Figure 1.1. In fact, now, in all higher education institutions, the content is the same for all students. This means that despite the knowledge and psychophysiological characteristics of students, they are taught from the first topic to the last. If there were an adaptive intellectual system for education, then the trainees could study only those topics that they do not know. That is, adaptive is the system where the initial knowledge of the user is considered, as well as the features of perception.

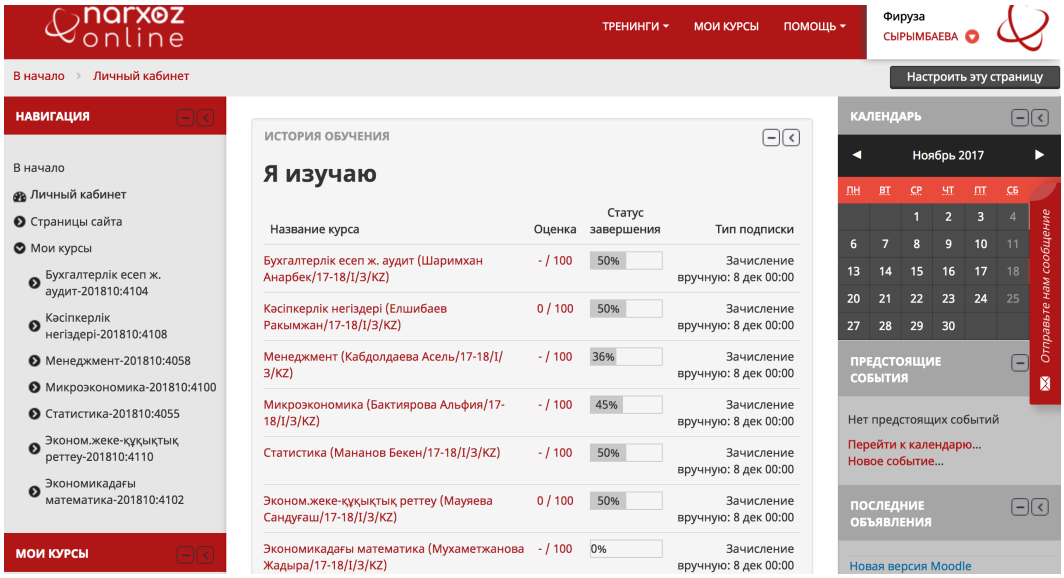

Figure 1.1 – Content for distance education

The competitve characteristics of information systems of distance education of universities of Kazakhstan are shown in Table 1.1.

Table 1.1 – Competitve characteristics of information systems of distance education of universities in Kazakhstan

№	Parameter	Astana IT University	IITU	Narxoz University	AlmaU	UIB	KazNU	Astana MU
1	Experience of successful implementation in educational institutions	implemented	implemented	implemented	implemented	implemented	implemented	implemented
2	The possibility of integrating the LMS with the electronic information environment of the customer	possible	possible	possible	possible	possible	possible	no
3	Open source	supports	supports	supports	supports	supports	supports	supports
4	Methodological and pedagogical features are taken into account	taken	taken	taken	taken	taken	taken	taken
5	Ability to use rich-media (fonts, pictures, videos)	possible	possible	possible	possible	possible	possible	possible

Continuation of the table 1.1

6	The presence in the system of means providing communication between the trainer and the students	present	present	present	present	present	present	present
7	The availability of tools for the customer to create new courses	present	present	present	present	present	present	present
8	Ability to import courses according to SCORM standards	possible	possible	possible	possible	possible	possible	possible
9	Ability to import courses according to IMS standards	possible	possible	possible	possible	possible	possible	possible
10	Structure	Modular	Modular	Modular	Modular	Modular	Modular	Modular
11	Web-oriented framework	present	present	present	present	present	present	present
12	Availability of technical support	yes	yes	yes	yes	yes	yes	yes
13	Multilingual interface	yes	yes	yes	yes	yes	yes	no

7 higher educational institutions of Kazakhstan were selected, such as: Astana IT University, International University of Information Technologies, Narkhoz University, AlmaU, UIB, Al-Farabi Kazakh National University and Astana Medical University. Since March 16, 2020, all universities have switched completely to distance education in connection with the pandemic worldwide. All of the above universities use the "Moodle" SDO, as there are advantages of the system: a modular web-oriented structure, support for multimedia data, it is possible to work with courses created according to international standards SCORM and IMS, as well as an open source system.

The most well-known foreign systems to:

1) WebCT (Developer: British Columbia State University, Canada) - implemented at the University of Texas (USA), McGill University (Canada);
2) Desire2Learn (developer: Desire2Learn, Canada) - implemented at the University of Minnesota (USA), the University of Iowa (USA), the University of Wisconsin (USA), etc.
3) ANGEL LMS (developer: ANGEL Learning, USA) - implemented at Indiana State University (USA), etc.;

4) Blackboard (developer: Blackboard Inc., USA) - implemented at Cornell University (USA), etc.;
5) Claroline (developer: CERDECAM Research Center, Belgium) - implemented in Universite Catholique de Louvain (Belgium), Haute Ecole Leonardo de Vinci (Belgium), Universidade de Vigo (Spain), etc.;
6) eFront (developer: EPIGNOSIS LTD, Greece) - implemented in the Ministry of Internal Affairs of the Republic of Poland, in state and educational institutions of the USA, Greece and Australia;
7) Metacoon (developer: metaVentis, Germany) - implemented at the University of Weimar (Germany), etc.;
8) Moodle (developer: MOODLE PTY, Australia) - implemented at the University of San Francisco (USA), etc.;
9) Sakai Project (developer: Sakai Foundation, USA) - implemented at the University of Michigan (USA), the Polytechnic University of Valencia (Spain), Stanford University (USA) and the Massachusetts Institute of Technology (USA) , etc.;
10) SharePoint LMS (developer: Microsoft, USA and Learningforce, Denmark) - implemented at Queen's University (UK), South Cheshire Collage (UK), etc.. [6]

The analysis of distance education revealed the following advantages: flexible schedule of the educational process, as well as the availability of materials and mobility. Disadvantages: self-discipline, responsibility, tolerance, commitment and perseverance.

1.2 Analysis of existing intellectual information and training systems

There are quite a few adaptive systems in the world, such as the Newton platform and the AltSchool school, shown in Table 1.2. There are 3 types of systems: a system for data collection, a system for output, and a personalization system. The data collection system collects information through testing or interviews. The system of conclusions, summarizes information and outputs data for the user, and the personalization system evaluates the student's capabilities and forms a learning strategy.

Table 1.2 – Review of existing adaptive systems

Functions	**Knewton platform**	**AltSchool**
Data collection system	collects information about the knowledge of students, the level of assimilation and understanding of the courses passed	interview, during which the Learner Profile
Terminal system	summarizes information based on the data obtained about the characteristics of the student, his	The application automatically tracks attendance, progress,

Continuation of the table 1.2

	reactions to changing the trajectory of learning	monitors the personal characteristics of the child and makes changes to his personal learning plan
Personalization system	on the basis of the data of the entire system, assesses the student's capabilities, and taking this into account, adjusts the goals and forms the optimal learning strategy for each student	Individual programs based on the needs and preferences of each child

Adaptive intelligent systems assume: flexibility of learning in an interactive educational environment, personalization and adaptation of learning, free access to content regardless of geography. Such technologies make it possible to develop educational and methodological materials, as well as to form individual learning trajectories.

1.3 Multicriteria learning model. Research of multicriterial decision-making model for educational information systems

Decision-making model. On the basis of multicriteria assessment model of knowledge can identify the main characteristics of the organization and control of automated learning process in information and training system. [7] These include:

- Knowledge level;
- The level of difficulty;
- The level of the reaction;
- The level of confidence.[8]

The level of knowledge - the level of current results to the user, based on the coefficient K2.[9]

Difficulty - fixed characteristics prescribed settings instructor.

The level of response - time assessment of the user's actions in response to any impact. The composition of response level includes coefficients K1, K3 and K5. [10]

The level of confidence - probability characteristic, is inversely proportional to the level of doubt. The structure includes a level of confidence coefficients K4 and K6. [11]

The decision on the basis of multiple criteria model in accordance with Figure 1.2 [12].

Decision-making mode information and training system based on the truth table for the four criteria in accordance with Table 1.3, 1 - max, 0 - min, 1 - knowledge, 2 - difficulty level 3 - the level of reaction, 4 - level of assurance [13].

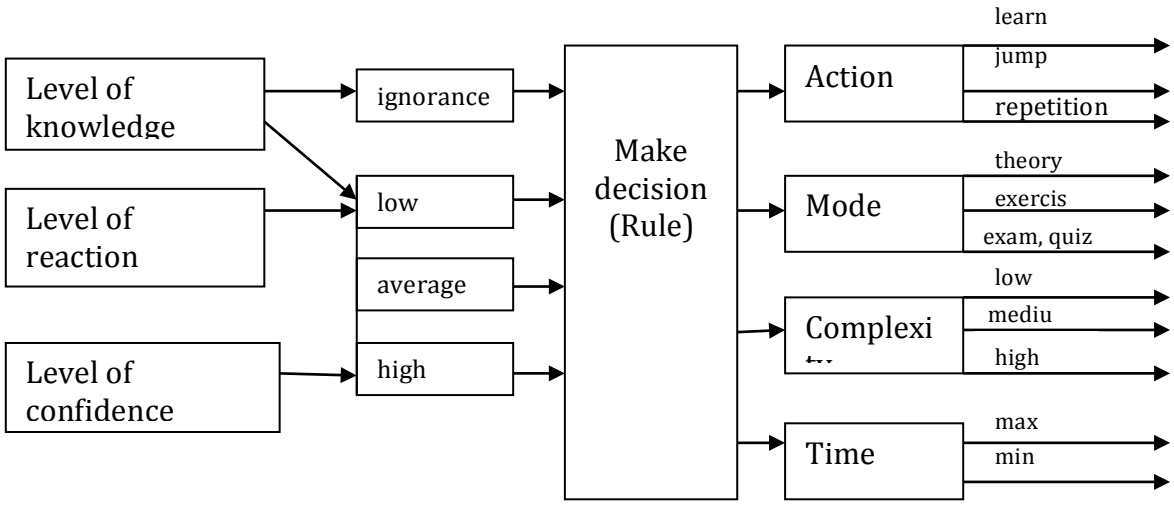

Figure 1.2 - Multicriterial decisions making model

Table 1.3 - The truth table of decision-making model on four criteria

State				Mode
1	**2**	**3**	**4**	
0	0	0	0	Training (opens. Mode)
0	0	0	1	Training (opens. Mode)
0	0	1	0	Training (opens. Mode)
0	0	1	1	Training (closed Mode)
0	1	0	0	Training (closed Mode)
0	1	0	1	Training (closed Mode)
0	1	1	0	Training (opens. Mode)
0	1	1	1	Training mode
1	0	0	0	Training mode
1	0	0	1	Adaptive mode
1	0	1	0	Adaptive mode
1	0	1	1	Setting a level of education
1	1	0	0	Adaptive mode
1	1	0	1	coaching mode
1	1	1	0	correction mode
1	1	1	1	mode self

Decision making educational element (action, mode of operation, complexity, time) on the basis of the current state of the educational element is achieved on the basis of the truth table of decision-making in accordance with Table 1.4. [14]

Table 1.4 - The current state of the educational element and decision

State			Decision of EE			
Knowledge	Reaction	Confidence	Complexity	Time	Action	Mode
Z0	RA	UA	A	max	jump	theory
Z0	RA	UB	A	max	learn	theory
Z0	RA	UC	A	max	learn	theory
Z0	RB	UA	A	-	learn	theory
Z0	RB	UB	A	-	learn	theory
Z0	RB	UC	A	-	learn	theory
Z0	RC	UA	A	min	learn	theory
Z0	RC	UB	A	min	learn	theory
Z0	RC	UC	A	min	learn	theory
ZA	RA	UA	A	max	repeat	exercise
ZA	RA	UB	B	max	repeat	-
ZA	RA	UC	B	max	jump	exam, quiz
ZA	RB	UA	A	-	repeat	exercise
ZA	RB	UB	B	-	repeat	-
ZA	RB	UC	B	-	jump	exam, quiz
ZA	RC	UA	A	min	repeat	exercise
ZA	RC	UB	B	min	repeat	-
ZA	RC	UC	B	min	jump	exam, quiz
ZB	RA	UA	A	max	repeat	exercise
ZB	RA	UB	B	max	repeat	-
ZB	RA	UC	C	max	jump	exam, quiz
ZB	RB	UA	A	-	repeat	exercise
ZB	RB	UB	B	-	repeat	-
ZB	RB	UC	C	-	jump	exam, quiz
ZB	RC	UA	A	min	repeat	exercise
ZB	RC	UB	B	min	repeat	-
ZB	RC	UC	C	min	jump	exam, quiz
ZC	RA	UA	B	max	repeat	exercise
ZC	RA	UB	B	max	repeat	-
ZC	RA	UC	C	max	jump	exam, quiz
ZC	RB	UA	B	-	repeat	exercise
ZC	RB	UB	B	-	repeat	-
ZC	RB	UC	C	-	jump	exam, quiz
ZC	RC	UA	B	min	repeat	exercise
ZC	RC	UB	B	min	repeat	-
ZC	RC	UC	C	min	jump	exam, quiz

The metric scale measuring the state of the educational element: first is ignorance (0-49%) measured by Z0. Second is low level of knowledge (50-74%) measured by ZA. Third is the average level of knowledge (75-89%) measured by ZB. Fourth is the high level of knowledge (90-100%) measured by ZC. Fifth is low level of reaction (0% -74%) measured by RA. Sixth is the average level of reaction (75-89%) measured by RB. Seventh is the high level of reaction (90-100%) measured by RC. Eighth is low level of confidence (0-74%) measured by UA. Next is the average level of confidence (75-89%) measured by UB. Then the high level of confidence (90-100%) measured by UC [15].

The organization of the learning process in information-learning system based on a measure of doubt for control need rules, which formed the knowledge base [16].

Decision-making model generate rules. [17] There are: if (REs knowledge - ignorance and reaction - low, average or high and confidence - low, average or high) THEN (complexity - low, the learning mode - the theory, the effect of educational elements - learn); if (REs knowledge - ignorance, low, average or high and reaction - low and confidence - low, average or high), time (time - MAX); if (REs knowledge - ignorance, low, average or high and the reaction - high and confidence - low, average or high), time (time - MIN); if (level of knowledge of UE - low. Medium or high and reaction - low, average or high and confidence - high) THEN (action educational element -Jump); if (knowledge UE – low, average or high and reaction - low, average or high and confidence - low or average) THEN (action educational element - Repeat); if (knowledge UE – low, average or high and reaction - low, average or high and confidence - Low) THEN (training mode - Exercise); if (knowledge UE – low, average or high and reaction - low, average or high and confidence - high) THEN (training mode – Exam, Quiz); if (knowledge UE - low or average and reaction - low, average or high and confidence - Low) THEN (difficulty - Low); if (knowledge UE - low and reaction - low, average or high and confidence - average or high) THEN (difficulty - average); if (the level of knowledge of UE - average and reaction - low, average or high and confidence - average) THEN (difficulty - average); if (the level of knowledge of UE - average or high and reaction - low, average or high and confidence - high) THEN (complexity - high); and last one is (knowledge UE - high and reaction - low, average or high and confidence - low or average) THEN (difficulty - average).

In the table 1.5 illustrates input data and decision-making model. If student has Educational element: learn, he/she should learn theory and spend minimum time and complexity must be low level [18,19].

Table 1.5 – decision-making model

Knowledge	Reaction	Confidence	Mode	Time	Complexity	Ed.element
78	80	70	exercise			repeat
50	70	77			average	
88	90	50	exercise	min	average	
99	49	66	exercise	max		repeat
7	49	44	theory		low	learn
67	66	5	exercise	max		repeat
77	78	44	exercise			repeat
88	90	40	exercise	min	average	
66	98	87		min		repeat
99	58	49	exercise	max		repeat
88	40	50	exercise	max		repeat
40	44	30	theory		low	learn
40	40	88	theory		low	learn
55	55	55	exercise	max		repeat
5	7	8	theory		low	learn
88	95	88		min	high	repeat
4	91	5	theory	min	low	learn
99	6	60	exercise	max		repeat
70	4	6	exercise	max		repeat
60	66	4	exercise	max		repeat
88	85	96	exam,quiz		high	jump

For example, fifth student has characteristics: knowledge is 7%, reaction is 49% and confidence is 44%. Model gives the following results: educational element is “learn”, mode “theory” and complexity is “low”.

In database one data exists with level of knowledge is 88%, level of reaction is 85% and level of confidence is 96%, which has educational element “jump”, mode: Exam or Quiz and complexity is high.

That student should pass Exam or Quiz for finishing course [20].

Decision-making model was created on the basis of measuring the level of doubt the user to control the learning process. The proposed idea makes it possible to reduce the probability of guessing the correct answer for a more objective assessment of knowledge and adapt the learning process on the basis of the knowledge base.

The results obtained in this study can be used for decision-making in the learning management information and education distance learning system. The practical value of the work lies in the fact that the use of information in the learning systems based on multicriterial decision making model obtained in the work leads to increase in the average score on the exam as compared to the control groups by about 0.3-0.4 points and reduce the amount of time required for learning about 1.5- 2.0 times.

On the basis of mathematical models and information obtained in the work created several computer applications: intelligent information and training system “Programming Languages Borland: Pascal & Delphi”, methodical complex “Mechanics. Molecular physics and thermodynamics” algorithmic learning system “camel”, information learning system for programming “Technology design software based on universal component in Delphi”, the interactive test suite of information technology. [21]

Distance education and e-learning have laid the foundation for a new global phenomenon: Smart Education, which is not so much about technology as it is about the new philosophy of education.

Developers of distance learning systems can not always fully take into account all the laws of the learning process, and therefore use not quite correct models. And even the most successful solutions in the field of online training rarely get widespread.

Undoubtedly, distance learning is by no means a complete substitute for traditional training, because it is difficult to replace live communication with a teacher or the educational atmosphere that develops in the classroom between subjects of the learning process. At the same time, modern ICT (information and communication technologies) can minimize the "narrow" places of distance learning. Modern pedagogy is also changing, nowadays transfer of part of the training load to the online mode is quite an acknowledged educational scenario. This can be done when developing educational content on a modern IT basis, for example, using MOOCs (a mass open educational course).

Today we can confidently state the existence of a new digital Z-generation of people for whom a mobile phone, a computer and the Internet are natural elements of their living space. In modern conditions, effective education is education without reference to time and place. This education teaches through everyday life. The transition to this mobile learning technology involves the application of new methods, approaches, principles of the learning process organization. When developing content for mobile devices, it is necessary to take into account that it is intended for young people of the Z-generation, as the method of preparing training content for them differs from the traditional learning content.

So, Smart technologies in education, such as mobile applications, are of great importance as they allow to optimize the costs of the university logistics, and also to raise the quality of educational services and products to a new level. It is Smart-technology that allows to develop revolutionary teaching materials, as well as to form individual trajectories of training.

In 2015, UNESCO published Recommendations on Mobile Learning Policies [22], which fully justifies the need for the introduction of mobile technologies in the educational process According to these recommendations, "In a world in which dependence on means of communication and access to information is growing, mobile devices will not be a transient phenomenon. As the capacity and capabilities of mobile devices are constantly growing, they can be used more widely as educational tools and take a central place, both in formal and informal education".

World experience shows that the competitiveness of the national economy as a whole depends on the development of information technologies. According to the World Economic Forum, the competitiveness index of the economies of states has a high level of correlation with the countries' information and communication technologies. According to the 2012 rating of the competitiveness of 142 countries of the world by the World Economic Forum, countries that are actively developing information technologies are ahead of Kazakhstan which ranks 51st in terms of creating demand for information technology (the USA - 13, Germany - 19, India - 63, Egypt - 96), and occupies the 104th place in the information technology business conditions place (USA - 21, Germany - 38, India - 72).

Scientific novelty consists in developing a method of objective measurement of students' knowledge on the basis of a multicriteria decision-making model.

The multicriteria model allows to objectively assess the student's knowledge level and, accordingly, to estimate the evaluation during computer testing. The model includes an original method of measuring the level of doubt, as a function of the time and complexity of issues differentiated by thematic blocks. And the assessment model itself being multicriteria, besides the function of doubt includes: the number of correct answers, the testing time, the use of additional programs, spent time on each response, the number of periodic passes in the testing process, and more. In diagnosing the level of knowledge based on computer testing, use of this method will reduce the probability of the influence of an accidental correct answer ("guessing") on the final result and to obtain the most objective assessment.

The proposed model includes the following factors:

- number of answers;
- the level of complexity of the questions depending on the time;
- spent on their response;
- time of all testing;
- doubt;
- number of passes;
- use of additional programs.

The practical significance of this model lies in the objectivity of assessing the level of knowledge, where computer testing is applied, especially relevant for the organization of distance learning using E & M Learning. Additional effects of the application of this model may be the reduction of the time for organization, intermediate control procedures, labor costs, operating costs, processing and analysis costs, and others, research planned in subsequent publications.

In conclusion, we can make a statement that the null hypothesis put forward by the authors is proven experimentally and the study demonstrates the original methodology for measuring the level of doubt of the tested in the information and training system. The use of this methodology allows making the most objective decision on the assessment of students' knowledge. In diagnosing the level of knowledge based on computer testing, the use of this methodology makes it possible to reduce the probability of the random factor influence, such as "guessing" on the final result, and to obtain the most objective assessment. Moreover, to control

the learning process, measuring the level of doubt allows taking into account the psychological characteristics of human behavior during training. [23]

1.4 Analysis of the identification of the need for the use of learning systems

A survey was conducted among 94 respondents in russian language about the use of an adaptive intelligent learning system. Questions were devoted to distance learning. Since the adaptive system in the first place can be used in distance learning. In one of the questions it was necessary to indicate from which city user filling out a questionnaire, or user would like to study. This will show the interest of residents of different cities. In accordance with Figure 1.3 showing cities, which respondent selected. Many of the respondents are from large cities, such as Almaty, Aktau, Shymkent and etc.

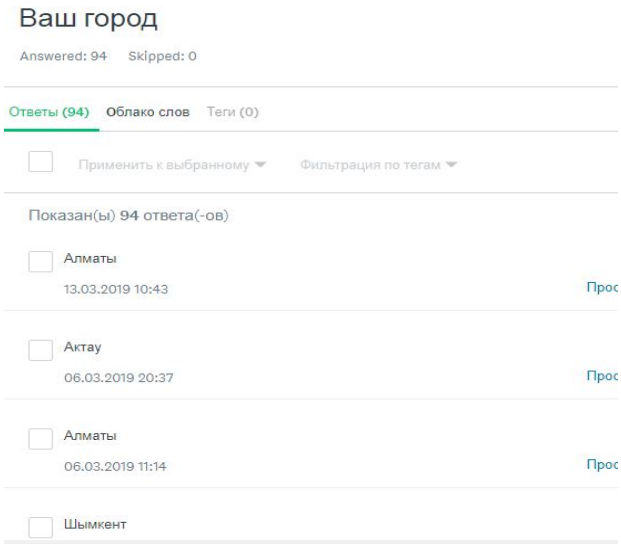

Figure 1.3 – Select respondent city

The following questions are intended to determine the social status (student, student, employee, retired, unemployed), on the use of distance learning. The question of the assimilation of information will also allow to determine the type of perception among users in addition to testing to determine the type of perception. In accordance with Figure 1.4 shown percentage of using distance learning. More than 60 percent did not use distance learning, 15 percent sometimes used it, more than 10 percent rarely, more than 5 percent often and 2 percent always used it. But at the moment, all students are using distance learning during the pandemic.

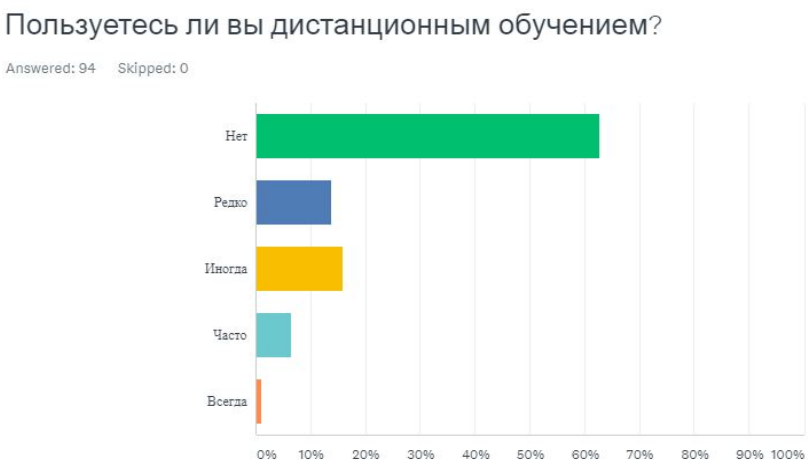

Figure 1.4 - Using distance learning

In accordance with Figure 1.5 shown learning information from respondents. More than 75 percent of respondents watch video lessons and easily assimilate information, 20 percent learn the material by reading lectures, and the remaining 5 percent learn the material by listening to audio materials.

Figure 1.5 – Learning information from users

In accordance with Figure 1.6 shows the answers to the question of a positive / negative attitude towards an adaptive intellectual system. Approximately 80 percent of respondents want to learn an adaptive intelligent system while sitting at home, and more than 20 percent of respondents do not want to learn such a system. [24]

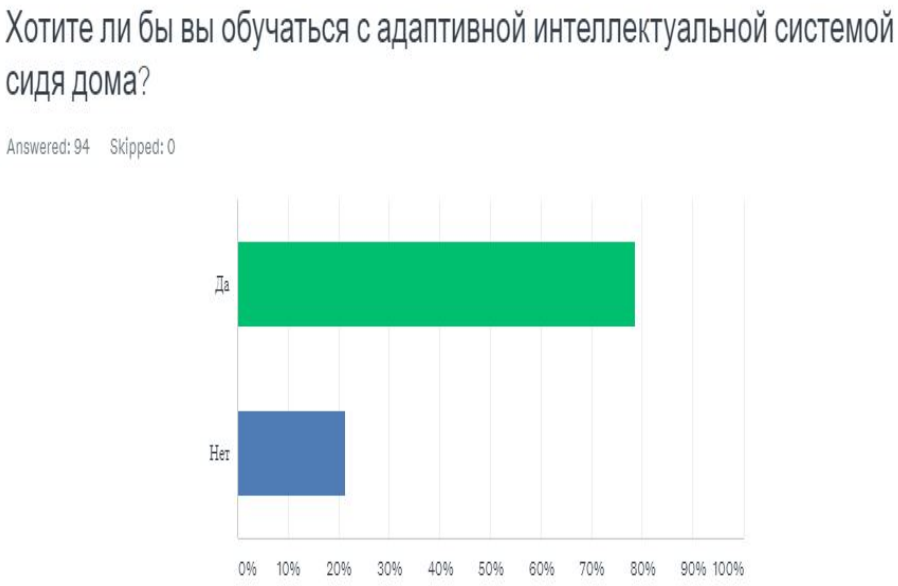

Figure 1.6 – Result of survey

The results of the questionnaire showed that respondents want to learn using an adaptive intellectual system and how relevant it is now.

Based on the results of the questionnaire, it can be concluded that students do not want to spend time on subjects that they already know, they also want an individual learning trajectory. The system must be adaptive intelligent for users.

1.5 Related works

Active learning in which an element of group discussion is incorporated is being attempted in classes at institutions of higher learning such as universities [25]. These efforts are aimed at using discussion to cultivate problem finding and problem solving skills and educate the future global workforce. e-Learning systems are vehicles that pro-mote such efforts [26]. The mode of traditional education is transforming with, for example, the rise of massive open online courses (MOOC) for which use has begun in the United States and JMOOC (Japan MOOC) [27, 28]. In this new mode, learners use an e-Learning system such as a MOOC to study content in advance, and then gather in a classroom to hold a group discussion to apply their accumulated knowledge. This type of class method is called a flipped classroom because the order of the class is the opposite that of a traditional class [29].

Toshiyuki Tojo, Osamu Ono, Norzaidah Binti Md Noh and Rubiyah Yusof [30] aim to implement the aforemen-tioned prelearning and group discussion in a consis-tent e-Learning system. We use the term collaborativee-Learning system to describe the feature that supportsgroup learning and to prevent confusion with traditionale-Learning. In this system, also use a humanoid robot tosupport learning. One project in the rapidly growing robotindustry is the research and development of human-robotinteraction for the purpose of creating a social, autonomousrobot capable

of conversing using various human-likemethods such as body language, hand gestures, facial ex-pressions, gaze, and touching.

Joksimovic. S., Poquet. O. , Kovanovic V. current gaps in the research, including a lack of solid frameworks to explain learning in open online setting. Finally, we put forward a novel framework suitable for open online contexts based on a well-established model of student engagement. Our model is intended to guide future work studying the association between contextual factors (i.e., demographic, classroom, and individual needs), student engagement (i.e., academic, behavioral, cognitive, and affective engagement metrics), and learning outcomes (i.e., academic, social, and affective). The proposed model affords further interstudy comparisons as well as comparative studies with more traditional education models. [31]

According to Firat M., Kilinc H., Yuzer T. motivation that initiates and sustains behaviour is one of the most significant components of learning in any environment. Accordingly, level of intrinsic motivation triggers and sustains the interest of the open and distance education students when it comes to learning on their own in e-learningenvironments. Despite a comprehensive literature regarding the motivation of those learning in traditional learning environments, the number of studies addressing the motivation of open and distance educationstudents in e-learning environments is not sufficient. In this context, this study aims at determining the level of intrinsic motivation of open and distance education students. Thus, data were collected from 1,639 distance education students in 22 programmes, through Intrinsic Motivation in e-Learning Questionnaire developed and validated to that end. Analyses carried out indicate that the level of intrinsic motivation of open and distance education students is high in e-learning environments, but there is not a statistically significant difference by gender, programme structure (graduate/undergraduate), instruction type (distance-blended), and academic disciplines. [32]

Muhammad Farhan, Sohail Jabbar, Muhammad Aslam, Mohammad Hammoudeh, Mudassar Ahmad, Shehzad Khalid, Murad Khan, Kijun Han wrote Students' interaction and collaboration using Internet of Things (IoT) based interoperable infrastructure is a convenient way. Measuring student attention is an essential part of educational assessment. As new learning styles develop, new tools and assessment methods are also needed. The focus of this paper is to develop IoT-based interaction framework and analysis of the student experience of electronic learning (eLearning). The learning behaviors of students attending remote video lectures are assessed by logging their behavior and analyzing the resulting multimedia data using machine learning algorithms. An attention-scoring algorithm, its workflow, and the mathematical formulation for the smart assessment of the student learning experience are established. This setup has a data collection module, which can be reproduced by implementing the algorithm in any modern programming language. Some faces, eyes, and status of eyes are extracted from video stream taken from a webcam using this module. The extracted information is saved in a dataset for further analysis. The analysis of the dataset produces interesting results for student learning assessments. Modern learning management systems can integrate

the developed tool to take student learning behaviors into account when assessing electronic learning strategies. [33]

Sean B. Eom and Nicholas J. Ashill argue that a significant reduction in dependent and independent variables and their measures is necessary for building an e-learning success model, and such a model should incorporate the interdependent (not independent) process nature of e-learning success. We applied structural equation modeling to empirically validate a comprehensive model of e-learning success at the university level. Our research advances existing literature on CSFs of e-learning and provides a basis for comparing existing research results as well as guiding future empirical research to build robust e-learning theories. A total of 372 valid unduplicated responses from students who have completed at least one online course at a university in the Midwestern United States were used to examine the structural model. Findings indicated that the e-learning success model satisfactorily explains and predicts the interdependency of six CSFs of e-learning systems (course design quality, instructor, motivation, student-student dialog, student-instructor dialog, and self-regulated learning) and perceived learning outcomes. [34]

Suman Bhattacharya, Samir Roy and Sankhayan Chowdhury presents an intelligent recognizer of the cognitive state of an e-learner as an integral part of confidence-based e-learning (CBeL) system. It addresses the problem of providing technology-driven pedagogical support to an e-Learner to achieve the desired cognitive state of mastery which is endowed by high levels of both knowledge and confidence. As per best of our knowledge concerned, no prior work has been done in the area of CBeL. The issue is crucial in the present scenario of teaching–learning in the twenty-first century where lifelong learning is gaining increasing importance vis-à-vis traditional classroom teaching–learning. However, self-learning is an indispensable mode of lifelong learning. It is felt that e-learning systems should have the capacity to simulate the behavior a human expert to identify the gap between the learners' cognitive state and the learning objective with the intension of guiding the self-learner take initiative to bridge the gap with appropriate action and eventually achieve his learning objective. An artificial neural network-based intelligent recognizer has been designed to identify the CBeL state of the learner on the basis of his performance in a CBeL test. This recognizer is the major agent that facilitates the implementation of the proposed CBeL system. Extensive experimentation has been carried out to ensure the performance of the recognizer. Results show ample evidence that the ANN-based intelligent recognizer is able to faithfully simulate the behavior of a human evaluator. [35]

Jose Marcelino, Maria; Pessoa, Teresa; Vieira, Celeste offered Learning Computational Thinking concepts by all has gained quite importance in last years. Scratch is also one of the most used products to promote it worldwide. Yet, most teachers face difficulties when using it in the classroom with their students. With this idea in mind we developed and run a distance education course under the Project of Distance Education of the University of Coimbra specially designed for elementary school teachers to learn both Computational Thinking concepts and Scratch via an e-learning course using Moodle as a Learning Management System. In this paper we

describe this first experience of teaching and learning online, the results obtained and also our future work about this course. Results of this first experience showed that it was possible for the trainees to learn Computational Thinking concepts and Scratch programming and also to develop useful products for their classroom practice using this modality of teaching and learning. [36]

Debiec, Piotr mentioned In the Internet era, students have increasingly lost interest in traditional lectures; as a consequence, their learning motivation and exam performance have decreased. The widespread adoption of learner-centered teaching methods that address this issue faces certain barriers, including: 1) the significant faculty effort necessary to prepare e-learning materials; 2) significant extra time required for active online communication with students; 3) student resistance to taking an active role in their education; and 4) lecturers' common belief that learner-centered teaching activities do not allow discussion of all the required topics. This paper presents a case study based on one offering of an introductory digital systems course taught with a combination of learner-centered strategies selected to overcome these barriers and improve student performance. These measures included: 1) improving the student-teacher relationship; 2) applying intriguing, inductive, and counterintuitive approaches to introducing new concepts; 3) adopting puzzle-based quizzes integrated with peer instruction; 4) using an audience response system; 5) replacing certain lectures with tutorials; 6) reducing course duration; and 7) using a graphics tablet. [37]

Gul, Sumeer; Mahajan, Iram; Shafiq, Huma identifies various issues and challenges faced by massive open online courses (MOOCs) while offering open online courses to vast number of learners. An exhaustive review of literature was taken up to carry out the study. The article reviewed the, issues and challenges faced by MOOCs. MOOC also provides an opportunity to groom the intellectual capacities of people at mass level. The initiative has removed all the barriers of time and space which is much evident in traditional educational system by offering courses at the doorstep of learners. To make universal education dream come true, there is a need to look at apprehensions which researchers have identified in the existing online learning environment. [38]

Farid, Shahid; Ahmad, Rodina; Alam, Mujahid propose a sustainable quality assessment approach (model) for the e-learning systems keeping software perspective under consideration. E-learning is becoming mainstream due to its accessibility, state-of-the-art learning, training ease and cost effectiveness. However, the poor quality of e-learning systems is one of the major causes of several failures reported. Moreover, this arena lacks well-defined quality assessment measures. Hence, it is quite difficult to measure the overall quality of an e-learning system effectively. [39]

Bradac, Vladimir; Walek, Bogdan presents a proposal, design and implementation of a new approach to adaptive e-learning systems. First, a proposal of a model is presented. This model aims at introducing adaptivity to current e-learning systems, which are rigid and limited in offering a truly personalised learning to individual students. Many of current e-learning systems enable personalised learning. However, in this paper, there is a new, innovative approach proposed for an adaptive

personalised e-learning system. The primary area of our research is English as a second language (ESL). Adaptivity in our view is considered as an ability of the system to adapt to student's knowledge and characteristics. This pedagogical perspective requires introduction of such processes that enable to work the pedagogical aspects of teaching/learning. The required processes are of informatics nature. [40]

Ren, Yi; Dai, Zhao-xia; Zhao, Xiao-huan present an on-line course applicability assessment (OCAA) to assist learners in course selection. This method is based on the statistical analysis of learners' individual characteristics and teaching strategy of specific on-line courses. Three main characteristics are taken into consideration, including 'learning style', 'learning behavioral type' and 'prior knowledge' which considerably affect e-learning effectiveness. Three on-line courses with different teaching strategies were adopted, and a two-step experiment was scheduled to establish the OCAA model and test its usability, usefulness and performance. The research findings show that the e-learning effectiveness is improved under the assistance of OCAA. Thus, OCAA could make it easier for learners to find suitable course matched with their own individual characteristics, and the e-learning effectiveness could be improved. [41]

Cohen, Anat focuses on the quantitative analysis of data accumulated on 362 students in three academic course website log files in the disciplines of mathematics and statistics, in order to examine whether student activity on course websites may assist in providing early identification of learner dropout from specific courses or from degree track studies in general. It was found in this study that identifying the changes in student activity during the course period could help in detecting at-risk learners in real time, before they actually drop out from the course. Data examination on a monthly basis throughout the semester can enable educators and institutions to flag students that have been identified as having unusual behavior, deviating from the course average. It was found that a large percentage of students (66%) who had been marked as at-risk actually did not finish their courses and/or degree. The presented analysis allows instructors to observe website student usage data during a course, and to locate students who are not using the system as expected. Furthermore, it could enable university decision makers to see the information on a campus level for initiating intervention programs. [42]

Advances and accessibility of Internet services around the world have transformed the traditional classroom learning into web-based e-learning systems. In recent years, designing adaptive e-learning systems has become one of striking topic of discussions in the literature. Additionally, integrating such systems with intelligent and adaptive systems that can measure the learning preferences of the user can enable learners to obtain the most suitable learning objects that might be matched with their learning styles. Moreover, even in the classroom teaching, knowing the learning styles of students can also help teachers to adopt appropriate learning materials for efficient learning. Hamada, Mohamed; Hassan, Mohamed described a case study on the integration of the learning style index into an adaptive and intelligent e-learning system [43]

Levina, Elena Y.; Masalimova, Alfiya R.; Kryukova, Nina I determined by the continuous information development of all spheres of education: integration of new knowledge, accessibility of information technologies and computer facility aids, professionalization and computerization of educational activities. They show the possibilities of using geo-information technologies in teaching outside the scope of their typical application (geographic, geodetic, geological education). The principles of designing the information environment for training on the basis of geo-information technologies are developed, which is built into the general information environment of higher education institute. The peculiarities of using geo-information technologies in non-core training are revealed, and the structure of learning environment modules based on geo-information technologies is developed and their content is described. [44]

Each learner has unique learning style in which one learns easily. It is aimed to individualize the learning experiences for each learner in e-learning. Therefore, it is important to diagnose complete learners' learning style and behaviour to provide suitable learning paths and automated personalized contents as per their choices. proposes some new dimensions of adaptivity like automatic and dynamic detection of learning styles and provides personalization accordingly. It has advantages in terms of precision and time spent. It is a literature-based approach in which a personalized adaptive learner model (PALM) was constructed. Sweta, Soni; Lal, Kanhaiya proposed learner model mines learner's navigational accesses data and finds learner's behavioural patterns which individualize each learner and provide personalization according to their learning styles in the learning process. Fuzzy cognitive maps and fuzzy inference system a soft computing techniques were introduced to implement PALM. Result shows that personalized adaptive e-learning system is better and promising than the non-adaptive in terms of benefits to the learners and improvement in overall learning process. Thus, providing adaptivity as per learner's needs is an important factor for enhancing the efficiency and effectiveness of the entire learning process. [45]

It has been several years since Massive Open Online Courses (MOOC) have entered the higher education environment and many forms have emerged from this new way of acquiring knowledge. Teachers have been incorporating MOOCs with more or less success in a traditional classroom setting to support various learning preferences, introduce this new way of learning to students, and to make learning available to those who might not be able to follow traditional instructions. Bralic, Antonia; Divjak, Blazenka blended learning model where a MOOC has been integrated in a traditional classroom. A learning outcomes based approach was implemented, that supported a balanced student workload. Qualitative approach was used to analyse students' learning diaries. Based on this research, benefits of integrating a MOOC with classroom based teaching were identified, as well as barriers that can hinder the successful implementation. [46]

In the work of Gregori, Pablo; Martinez, Vicente; Jose Moyano-Fernandez, Julio, aim is to design and analyze the implementation of a number of guidelines that allow us to effectively unify a high-quality teaching methodology and the use of new

technologies in distance learning. The intention is to guide the relationship between teacher and student, establishing a clear, coordinated and evaluable method of action to make any changes that might be needed to improve the learning experience. We will focus on the analysis of the dropout rate to establish patterns of behavior at all levels that allow us to decrease this indicator. [47]

The ECO European project (E-learning, communication, Open-Data) has contributed to dissemination of the MOOC throughout the European continent, therefore, proving the expansive development of these models of education that allow participation and interaction through social networks in the platform of the course. The special feature about this project is the student's empowerment, which has been publicized as a proposal that allows the users to create their own sMOOC. Gil Quintana, Javier; Martinez Perez, Jorge have analyzed this massive, open and online model course, describing the advantages and disadvantages of them, besides de formative and motivating factors that guide the students, not only to enroll in particular courses but also to request for their own space to be able to create a sMOOC. [48]

At present, the practice development of massive open online Courses (MOOC) is prior to the theoretical research about MOOC in academic circles. Through document analysis of Tang, Shujun, participant observation and case analysis of MOOC, we find that the connotation of MOOC is mainly annotated from three dimensions: curriculum form, education model and knowledge innovation. According to the learning theory and teaching practice of MOOC, the teaching mode of MOOC is divided into three kinds: MOOC based on content, MOOC based on network and MOOC based on task. Compared with the traditional courses, MOOC has intrinsic characteristics such as large scale, openness, networking, personalized and participation, which includes the online learning effectiveness, the mastery learning, the interactive cooperation and the learning mechanism of complex system self-organization core. Through the analysis of MOOC learning mechanism, goal, characteristics and advantages, according to the investigation on the use of MOOC at home and abroad, combined with the characteristics of computing discipline, the two aspects of MOOC teaching function: one is to use the learning records on MOOC to study pedagogy, and create new educational theory; another is to use MOOC for the flipped classroom. [49]

For the first time, a multicriteria method for measuring the level of knowledge of students during computer testing based on the level of doubt for E & M-learning has been described by the authors in a number of works, as discussed by Serbin V., Smaiyl A., Tolebayeva K., Gorbunov Y., [7, 50, 51]) and applied for distance training.

1.6 Research objectives

The review and analysis showed that the problem with the content sequence exists and has not been solved, in this regard, this paper examines the features of the

relationship of differences between matrials, determines the content sequence based on semantic data analysis, and also implements a system with an individual learning process.

The following main tasks are solved in the work:

1) Review and analysis of existing information systems for educational resources;
2) Models and methods for building an educational resource;
    2.1) Formation of the content of the educational resource;
    2.2) Compilation of content based on semantic analysis;
    2.3) Determining the sequence of content;
    2.4) Management of the number of training elements in accordance with Microlearning;
3) Development of the system in the form of a web portal;

Conclusions on the first section

This section provides an overview and analysis of existing adaptive information systems for education. In traditional training, there are shortcomings in the formalization of processes, since the training is the same for all students. The content is the same, and the types of perception of users are different, it is also possible that the user knows the entire course, but he must teach it. The traditional training system is not flexible, not individual for each user. The existing methods of formalization of training are weak, and it is desirable to increase the effectiveness of training by individualizing training. The development of the system using methods and models will increase the effectiveness of training.

The section also mentions a multi-criteria learning model, which also gives good learning efficiency.

The literature review showed that there is a trend of individual training in the world. Many scientists have developed adaptive systems for users.

An adaptive intelligent system will help all users to effectively spend time on training, while gaining the necessary knowledge.

2 METHODOLOGY FOR CREATING AN INTELLIGENT INFORMATION AND TRAINING SYSTEM

In this section of the dissertation, the methodology of creating an intelligent information and training system is considered. In particular, the content of the learning process should depend on the user's perceptual modality. And semantic content analysis with the help of ontology will provide an individual trajectory for the student.

2.1 Content and type of perception

In connection with the informatization of education, ideas of adaptive learning were also used in computer training. The main requirement that the information training system should be designed to take into account the principles of adaptive learning is to ensure the learning process (both within the university and in distance education) in accordance with the individual characteristics of the learner. This task can be solved by implementing in the training system various techniques and methods related to different variants of the training system functionality and various ways of its implementation.

Content, more precisely, the material is divided into 1) difficulty (easy, medium and difficult level); 2) the sequence (A, B, C); 3) structure; 4) content; 5) the size of the form (duration); 6) the form of the material (text, video, sound-audio, interactive, slide) 7) a combination (complex) 8) the formation of a unique text material on the data array upon request. [52]

Information properties are expressed:

1) the content of educational information. Factual information, concepts, skills and abilities may be the subject of memorization. Teaching methods, its character depends to a decisive extent on what type of educational material is the subject of memorization.

2) Form of educational information. The effectiveness of the chosen form of presentation depends on whether it corresponds to the nature of the educational material and how it corresponds to the thinking of the students.

3) The complexity of educational information that affects the efficiency of learning, its speed and accuracy. The difficulty for the user to memorize a new educational material, other things being equal, depends on its connection with the knowledge and skills that he already has, as well as with experience in using them. And the more there are such connections, the information is easier to learn.

4) Significance, importance of educational information. Some information or actions may be important in their own right or for mastering the subsequent material. Some educational information may carry educational, practical (business), ethical (moral), aesthetic (artistic), social (social), and receptive (pedagogical) significance.

5) meaningfulness. On the basis of observations and experiments, it has been established that the material that is meaningful is memorized faster, it is reproduced with a smaller number of errors, and is stored longer and more completely.

6) Structural. Learning any material makes it easier as the structure increases, i.e. logical, semantic and syntactic links of its parts. The more links there are with the old, the closer the links of each subsequent part to the previous, the easier it is to learn, therefore one of the principles of learning is the principle of consistency and systematicity. Difficulties arise when the links in factual and descriptive information are not clearly distinguished and masked. The arrangement of pieces of instructional information also plays a role.

7) The amount of educational information influencing the learning of educational material is its volume, i.e. the number of its individual elements. Measuring the number of elements of educational material is not enough, because the student learns not what is written in the textbook, but what he gets as a result of the mental processing of the text and the expression of these results in terms of his experience. The amount of meaningful educational information is measured only indirectly by the number of new concepts that need to be learned, the relationships that it establishes, or the number of suspicions it contains.

8) Emotionality. The emotional features of educational information is the attractiveness of the material for the user, his ability to evoke certain feelings and experiences in a person. Studies show that material that causes strong positive feelings is easier to learn than boring and indifferent.

Classification of lectures by forms can help students not to waste time.

1 Problem lecture. The task of the teacher: to create a problem situation, to encourage students to search for solutions to the problem, step by step bringing them to the desired goal. For this, a new theoretical material is presented in the form of a problem problem. In her condition there are contradictions that must be discovered and resolved.

2 Lecture - visualization. The visualized lecture is a verbal information transformed into a visual form. Depending on the educational material, various forms of clarity are used: 1) natural (minerals, reagents, machine parts); 2) figurative (slides, drawings, photos); 3) symbolic (schemes, tables). [52]

Psychophysiological abilities by type of thinking are divided into visuals, audials and kinesthetics. Diagnostics of the dominant perceptual modality of S.Efremtsev is a technique of the leading channel of perception. Visitors perceive information through images, memory, imagination. Visuals have increased gesticulation in accordance with Figure 2.1. Audials perceive information through hearing in accordance with Figure 2.2. It is possible to identify audials by asking a simple question: what did he do yesterday? If he looks to the side and to the left, he remembers truthful information; if he looks to the side and to the right, he comes up. Kinesthetic perceive information through feelings, emotions. It should be noted that a person completely belonging to one type of thinking does not exist. In each of us there is a piece from each of them, but at the same time we have a dominant type of perception. According to statistics, in 35% of people the dominant type of perception is visual, in 25% of people the dominant type of perception is audial, and in 40% of people the dominant type of perception is kinesthetics. [54]

Depending on the characteristics of perception and processing of information, users can be divided into 2 categories: audials and visuals.

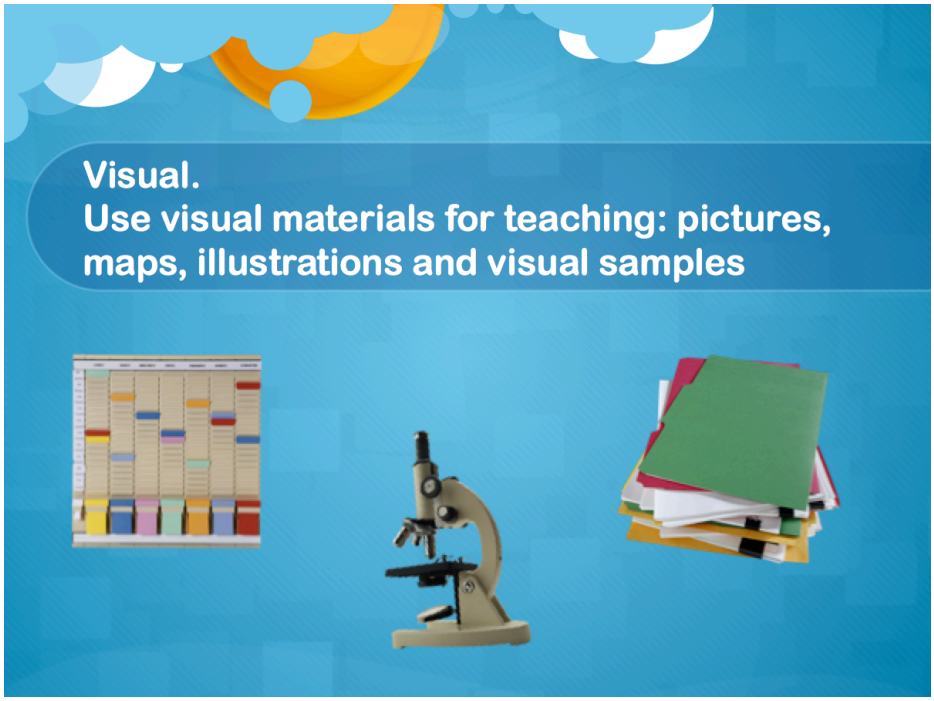

Figure 2.1 – Content for visuals

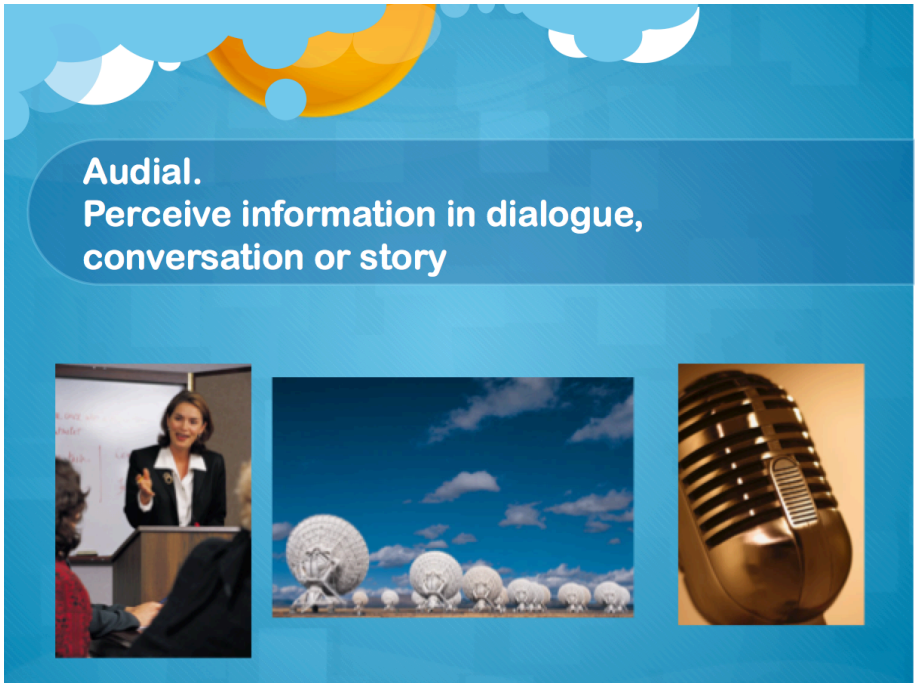

Figure 2.2 - Content for audials

For visuals, techniques for visualizing content: diagrams, graphs, illustrations, demonstration models, experiments and experiments.

For audials: Lecture material, audio courses, video, audio books.

In the course “Database in Information Systems” at the International University of Information Technologies, a test was conducted to determine the type

of thinking for 3 groups of students according to the method of the leading perception channel. The list of students is shown in Table 2.1, where perception types are divided by columns. This study did not consider the type of perception - kinesthetics. According to the test result, a double perception type prevails in some students: auditory and visual.

Table 2.1 – Separation of students by type of perception

N	Name	Audial	Visual	Audial and visual
1	Student_1	1	0	0
2	Student_2	0	1	0
3	Student_3	1	0	0
4	Student_4	0	1	0
5	Student_5	1	0	0
6	Student_6	0	1	0
7	Student_7	0	1	0
8	Student_8	1	0	0
9	Student_9	0	0	0
10	Student_10	0	1	0
11	Student_11	0	0	0
12	Student_12	1	0	0
13	Student_13	1	0	0
14	Student_14	0	1	0
15	Student_15	1	0	0
16	Student_16	1	0	0
17	Student_17	1	0	0
18	Student_18	0	1	0
19	Student_19	0	1	0
20	Student_20	1	1	1
21	Student_21	1	0	0
22	Student_22	1	1	1
23	Student_23	1	1	1
24	Student_24	0	1	0
25	Student_25	1	0	0
26	Student_26	1	0	0
27	Student_27	1	0	0
28	Student_28	1	0	0
29	Student_29	0	1	0
30	Student_30	0	1	0
31	Student_31	1	0	0
32	Student_32	0	0	0
33	Student_33	0	1	0

Continuation of the table 2.1

34	Student_34	1	1	1
35	Student_35	0	1	0
36	Student_36	1	0	0
37	Student_37	0	1	0
38	Student_38	1	1	1
39	Student_39	1	0	0
40	Student_40	1	0	0
41	Student_41	1	0	0
42	Student_42	1	0	0
43	Student_43	1	1	1
44	Student_44	1	0	0
45	Student_45	1	1	1
46	Student_46	0	0	0
47	Student_47	1	1	1

Data visualization in R studio. The system of statistical analysis and visualization of data R consists of the following main parts: a high-level programming language R, allowing one line to implement various operations with objects, vectors, matrices, lists, etc.; a large set of data processing functions, collected in separate packages (package); a developed support system that includes updating the components of the environment, interactive assistance and various educational resources intended both for the initial study of R and subsequent consultations on the difficulties encountered. [55]

Direct training in practical work in R consists of a) mastering the structures of the R language and familiarity with the features of calling functions that perform data analysis, and b) acquiring skills in working with programs that implement specific methods of data analysis and visualization.

For example, R Studio's free integrated development environment is an excellent option.

To read Table 1, need to export to a csv file, and write the following in R Studio: *data <- read.csv2("/Users/apple/Desktop/for research.csv")* in accordance with Figure 2.3. There is also a command summary (data) for viewing the full description of the file.

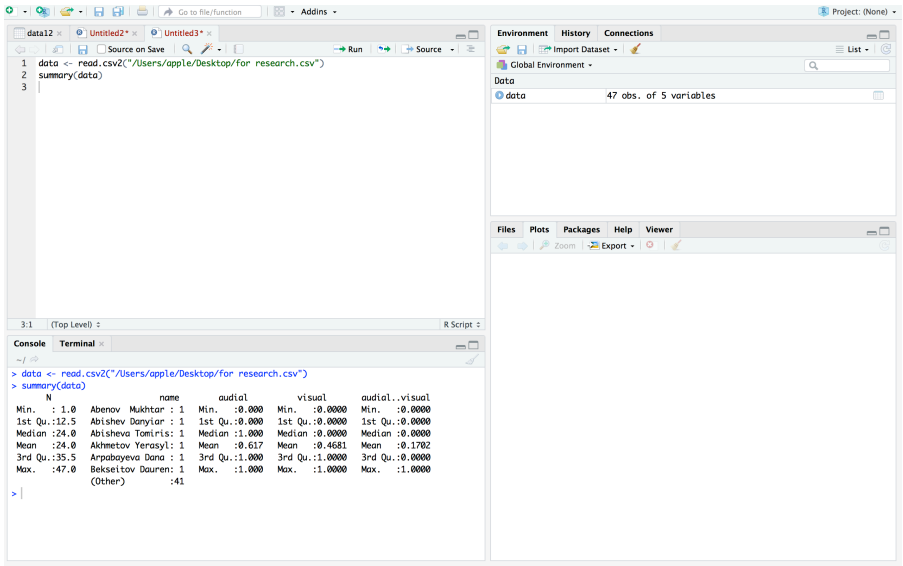

Figure 2.3 – Read file in R Studio

Commands like plot(ecdf(data$audial)) and hist(data$audial) to get data visualization. In accordance with Figure 2.4 shows a histogram for students-audial.

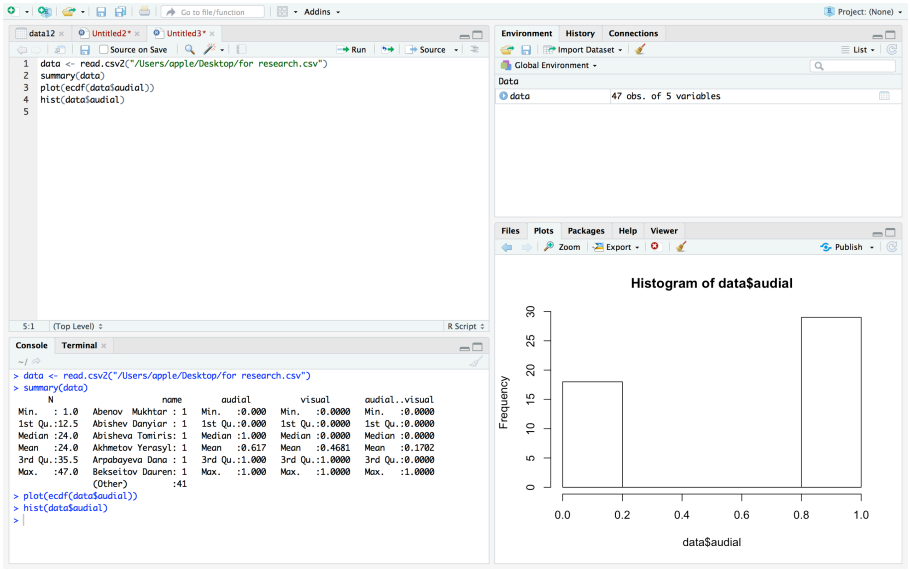

Figure 2.4 – Histogram for audial students

It is also possible to visualize data in the form of boxplot, or rather data about audials, in accordance with Figure 2.5.

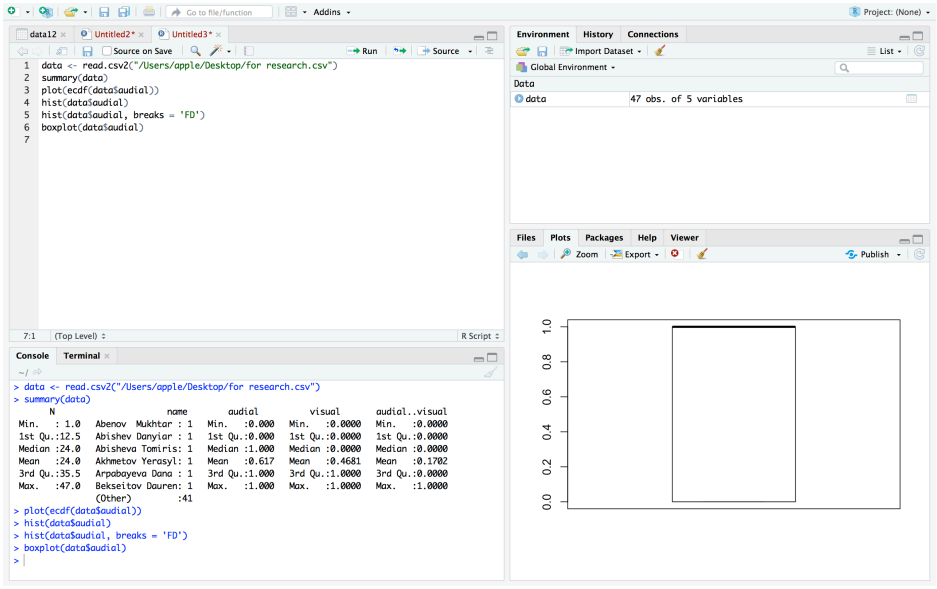

Figure 2.5 – boxplot of audial data

In R studio, data appeared, totaling 47 students. Audials 29 students. Double perception type: audials and visuals, a total of 8 students in accordance with Figure 2.6.

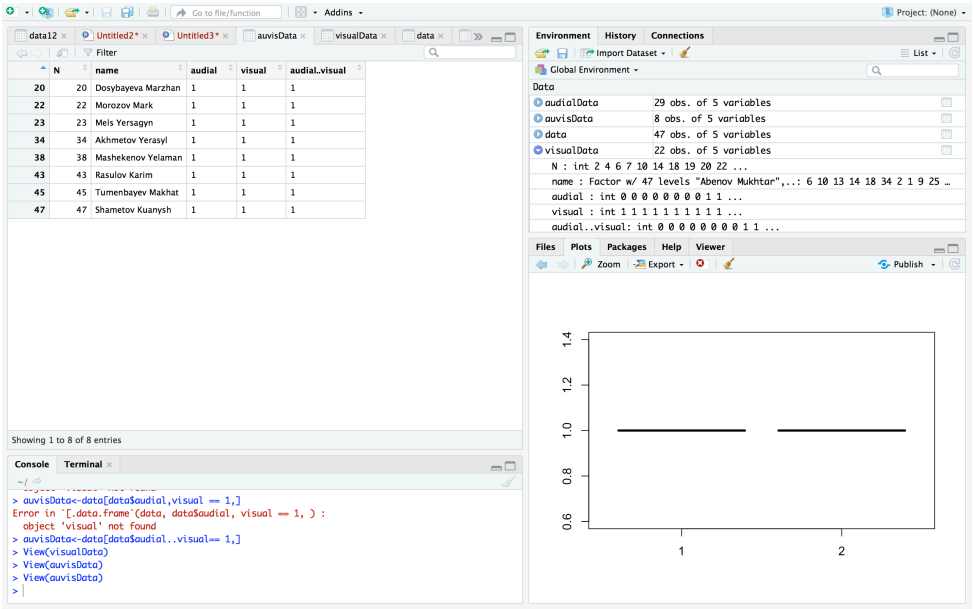

Figure 2.6 – Double perception type: audials and visuals

Adaptive information and training system will help improve the effectiveness of learning through the methodology of the leading channel of perception. Analysis of the use of an adaptive system allows to draw conclusions:

- Such classes creates support for thinking, develops visual modeling skills, which is a way to increase not only the intellectual, but also the professional potential of students.

- The choice of ways to achieve and types of visibility depends on the topic. Guided by the principles of feasible difficulties, in the presentation of topics difficult for perception and understanding, containing a large amount of concentrated information, it is advisable to use a combination of graphic and symbolic visibility. For example, the scheme is universal, but rather difficult for perception means of visualization. Therefore, it is recommended to carry out on the basis of the figure.

- The use of classes should be based on consideration of the psycho-physiological capabilities of students, their level of education and professional affiliation. [56]

During the testing of the information system, 94 2nd-year students participated. The distribution by type of perception in accordance with Figure 2.7.

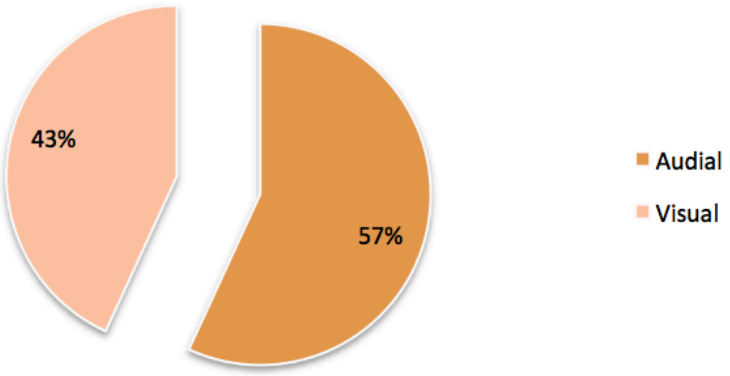

Figure 2.7 – Distribution by type of perception

Visualization helps to make the right data analysis for an adaptive learning system.

2.2 Methods of data processing in educational systems

The education system has always generated a significant amount of data. The only question was how to start working with these data at the system level: analyze them, make decisions based on them. [9]

In the field of education, there are five main types of data:

1) Personal Information;
2) Data on the interaction of students with electronic learning systems (electronic textbooks, online courses);
3) Data on the effectiveness of teaching materials;
4) Administrative (system-wide) data;
5) Forecast data; [10]

Big data is a term that describes the large volume of data – both structured and unstructured – that inundates a business on a day-to-day basis. But it's not the

amount of data that's important. It's what organizations do with the data that matters. Big data can be analyzed for insights that lead to better decisions and strategic business moves.

While the term "big data" is relatively new, the act of gathering and storing large amounts of information for eventual analysis is ages old. The concept gained momentum in the early 2000s when industry analyst Doug Laney articulated the now-mainstream definition of big data as the three Vs:

Volume. Organizations collect data from a variety of sources, including business transactions, social media and information from sensor or machine-to-machine data. In the past, storing it would've been a problem – but new technologies (such as Hadoop) have eased the burden.

Velocity. Data streams in at an unprecedented speed and must be dealt with in a timely manner. RFID tags, sensors and smart metering are driving the need to deal with torrents of data in near-real time.

Variety. Data comes in all types of formats – from structured, numeric data in traditional databases to unstructured text documents, email, video, audio, stock ticker data and financial transactions.

At SAS, we consider two additional dimensions when it comes to big data:

Variability. In addition to the increasing velocities and varieties of data, data flows can be highly inconsistent with periodic peaks. Is something trending in social media? Daily, seasonal and event-triggered peak data loads can be challenging to manage. Even more so with unstructured data.

Complexity. Today's data comes from multiple sources, which makes it difficult to link, match, cleanse and transform data across systems. However, it's necessary to connect and correlate relationships, hierarchies and multiple data linkages or your data can quickly spiral out of control. [11]

Modern technologies allow you to collect and analyze large amounts of information. "Big Data" in the field of education is becoming more and more saturated with every day. It is associated with the possibility of a significant transformation of the educational process and the pedagogical technology itself, which, due to constant improvements, can be brought to a high degree of perfection. Previously, no one collected such data, because there would not have been enough resources to record, store, and even more so for analysis. [3, 12]

Kazakhstan in case to compare with other development countries has very good e-government infrastructure where main player is JCS NIT www.nitec.kz – National Information Technologies, Government Company that provide almost all data bases (structured) such as data base of all citizens, data base of all business (companies, entrepreneurships, and other all types of organizations), data of all addresses and real estate. There are two main principles of above databases are personal identification number and business identification number. Both identification numbers are fully integrated, through e-government GATE with other databases such as; police, court, procurator, tax committee, social and other government databases, it shown in Fig.2. Moreover, public and private organizations also integrated through e-government GATE, there are banks, insurance, education,

transport etc., so we can identify that Kazakhstan almost ready to use big data technologies. However, there are many databases those not ready or not exist by unknown reasons, few of them are healthcare are statistics. [57]

2.3 Intellectual system for education

Intellectual system for education can help users to learn effectively. The information system assumes:

- flexibility of learning in an interactive educational environment;
- personalization and adaptation of training;
- diverse content by type of perception;
- free access to content regardless of geography.

It is the intellectual system that allows the development of revolutionary teaching materials, as well as the formation of individual learning paths.

System adaptability. The adaptability of the system is that each user will have their own learning path. When entering the system, users pass 2 tests that will determine the type of perception and level of knowledge for a specific course.

Intelligent information systems are a natural result of the development of conventional information systems. They have concentrated the most high-tech technologies with a high level of automation not only of the processes of preparing information for decision-making, but also of the processes of developing solutions based on the data obtained by the information system.

Intelligent information systems are a natural result of the development of conventional information systems, they have concentrated the most knowledge-intensive technologies with a high level of automation not only of the processes of preparing information for decision-making, but also of the processes of developing solutions based on the data obtained by the information system.

To solve problems, an AIS must have certain properties:

- solve problems described only in terms of soft models, when the dependencies between the main indicators are not completely defined or even unknown within a certain class;
- the ability to work with uncertain or dynamic data that changes during processing allows the use of AIS in conditions where data processing methods can change and be refined as new data becomes available;
- the ability to develop the system and extract knowledge from the accumulated experience of specific situations, which increases the mobility and flexibility of the system, allows it to quickly master new areas of application;
- the ability to use information that is not explicitly stored, but is derived from the data available in the database, allows you to reduce the volume of stored factual information while preserving the wealth of information available to the user.

The distinctive features of the IC in comparison with conventional ICS are as follows:

- interface with the user in a natural language using business concepts specific to the user's domain;
- presentation of the model of an economic object and its environment in the form of a knowledge base and means of deductive and plausible conclusions, combined with the ability to work with incomplete or inaccurate information;
- AIS solutions have "transparency", i.e. they can be explained to the user at a qualitative level;
- expert systems are able to supplement their knowledge in the course of interaction with an expert;
- the ability to automatically detect business patterns in previously accumulated facts and include them in the knowledge base(so-called machine learning).
- The AIS gives the user a" ready-made " solution that is not inferior in quality and efficiency to the solution of a human expert; the use of a specific component – the knowledge base [58].

2.4 Creating an ontology for a training system

Semantic content analysis using ontology

- Ontology
- Ontology Editor Protégé
- Ontology development includes:
- definition of classes in ontology;
- the arrangement of classes in a taxonomic hierarchy (subclass-superclass);
- definition of slots and description of the allowed values of these slots;
- filling in instance slot values.

Ontology. Ontology is the embodiment of conceptual knowledge about the subject area consists of the following structural components: taxonomy; description of the relations in which the objects of the subject area are located.

Using special technical tools - the Protégé ontology editor, OntoEdit.

The description of the domain ontology should be presented in the OWL Lite language.

Protégé – a freely distributed Java program designed for building (creating, editing, and viewing) ontologies of a particular application area. It includes an ontology editor that allows you to design ontologies by deploying a hierarchical structure of abstract and concrete classes and slots. Based on the generated ontologies, Protégé allows you to generate knowledge acquisition forms for introducing instances of classes and subclasses. This tool supports the use of the OWL language and allows you to generate HTML documents that display the structure of ontologies. Since it uses the OKBC framework model of knowledge representation (OKBC-Open Knowledge Base Connectivity), it allows you to adapt it for editing models of subject areas that are not represented in OWL, but in other

formats (UML, XML, RDF/RDFS (Resource Description Framework), DAML+OIL, etc.).

The development of the ontology includes:

1) defining classes in ontology;

2) the arrangement of classes in a taxonomic hierarchy (subclass-superclass);

3) definition of slots and description of the allowed values of these slots;

4) filling in the values of the instance slots.

After that, we can create a knowledge base by defining individual instances of these classes, entering a value in a specific slot, and additional restrictions for the slot.

Use the ontology library: 1. Ontolingua; 2. DAML; 3. UNSPSC; 4. RosettaNet; 5. DMOZ.

Taxonomy levels: Upper level, Middle level, and Lower level.

Slot is properties associated with classes (attributes). Slots can have facets that describe the type of value, the allowed values, the number of values (power), and other properties of the values that the slot can accept. A string is the simplest type of value that is used as a simple string. The number is float/integer. Boolean slots – true/false. Numbered Slots-defines a list of specific allowed values. Slots-ins.

In general, several types of object properties can become slots in the ontology:

- internal properties, such as the taste of wine;

- external properties, such as the name of the wine and the area in which it was produced;

- parts, if the object has a structure, they can be both physical and abstract " frequent»;

- relations with other individual concepts, this is the relationship between individual members of the class and other elements (example, wine producer, grapes)

2.4.1 Defining slot facets

The allowed classes for instance-type slots are often referred to as the slot value range. The classes to which the slot is bound, or the classes whose property the slot describes, are called the slot domain. When defining a domain or range of slot values, find the most common classes or class that can be a domain or range of slot values, respectively. On the other hand, don't define a domain and range of values too general: all classes in the slot domain should be described by the slot, and instances of all classes in the slot value range should be potential slot placeholders..

2.4.2 Creating instances

The last step is to create separate instances of the classes in the hierarchy. To define a separate instance of a class, you need:

1) select a class;

2) create a separate instance of this class;

3) enter the slot values.

Organization of metadata in a data warehouse Metadata is structured, encoded data that describes the characteristics of media objects that contribute to the identification, detection, evaluation, and management of these objects.

Types of metadata:

1) Source system metadata: specifications of data sources, such as repositories; descriptive information (update frequency, legal restrictions, and access methods); process information, such as task schedules and extraction codes.
2) Data transformation metadata: information about data acquisition (planning, data and results transfers, and file usage information; managing dimension tables, such as defining changes and assigning surrogate keys; transformation and aggregation, for example, data expansion and display, DBMS loading programs (scripts), data aggregate definitions; documentation of checks, operations, and logs, for example, data transformation logs and data origin tracking records).
3) DBMS metadata: the contents of the DBMS system tables; recommendations for processing.

Several levels of description are used for metadata, shown in Table 2.2.

Table 2.2 - Description of metadata levels

Level Description	**Level Description**
Application layer (external data sources)	Describes the structure of data in operational databases and other data sources. Usually, this level is quite difficult for an untrained user to understand and is application-oriented
Data Warehouse Core Layer	Describes the structure and relationships of data in a Data Warehouse
End user level	Describes the data structure in the Data Warehouse in terms of the end-user domain

Metadata can be used: actively, by storing specific semantic aspects.

2.4.3 Methods of semantic content analysis

The method of component analysis in determining the semantics of each word in the text. Then the meaning of the whole sentence is the set of meanings of all its words.

The method of semantic cases. Based on component analysis, it describes the semantics of a sentence, identifying semantic units. Center-predicate, Agent – target object (no clear criteria for identifying semantic roles).

The method of thematic classification based on the model of the structural representation of the text. It consists in identifying a set of key elements that are often repeated in the text, and then considering them in context by analyzing the structural representation of the text.

The associative experiment is used to make an assessment of the emotional coloring of texts.

Semantic differential. The essence of the method is to obtain an estimate of different concepts when using scaling. It is used to select the most appropriate synonyms from several. For example, performance evaluations: fast/long, predictable/unexpected, regular/random.

Latent semantic analysis. It consists in identifying the meanings of words (in context) using statistical document processing (semantic proximity of words).

Content analysis is the study of information and its translation into quantitative indicators for further statistics. Studies documents in their social context.

Thesauruses. The semantic relations between the concepts of the same subject area are described. A special type of dictionary that specifies the semantic relations between lexical units.

Ontology is mapping the relationships of objects, a specific subject area. An ontograph is a representation of an ontology in the form of a directed graph.

2.4.4. Semantic content analysis using ontology

The ontology includes sets of concepts, multiple relationships between concepts and functions for interpreting concepts from a set of concepts and a relation between concepts (1).

$$Q = <C, R, F> \quad (1)$$

Q – ontology;
C – sets of concepts;
R – multiple relationships between concepts;
F - functions for interpreting concepts from a set of C and a relation from R.

In Protégé, the ontology structure consists of classes, properties between classes or instances (Object Properties), properties between instances and data (Data properties), from instances (Individuals), and axioms.

Creating properties. Concepts can be linked by various kinds of relationships that link classes together and describe them. Domains specify the main concept, and Ranges specifies the object that describes it. InverseOf denotes the opposite property of this object.

Instantiation. Instances are separate representatives of a class of entities or phenomena. An instance is the lowest element of the chain, a specific detailed object of some kind that can be linked to others or to data.

The creation of axioms. The axioms specify the terms of the correlation of classes and relationships, expressing obvious approval. An axiom can be understood as a statement that is entered into the ontology in a ready-made form, from which

other statements can be derived. Axioms allow you to make inferences within the ontology, define rules that allow you to automatically add information to the ontology, as well as restrictions imposed on any relations. In Protégé, axioms can be described using the EquivalentTo or SubclassOf function. Equivalent classes are intended to describe a class using others that would then replace them [59].

An ontograph is a representation of ontology as a directed graph in accordance with Figure 2.8.

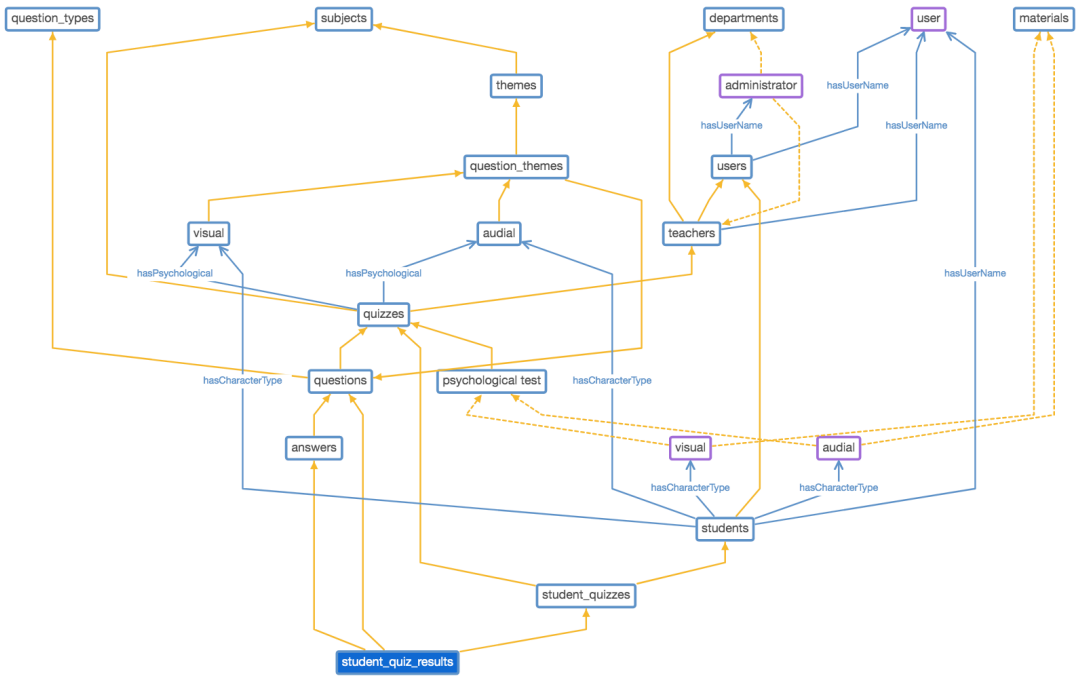

Figure 2.8 – ontograph of information system for education

Defining classes in the ontology. The following 16 classes with attributes in accordance with Figure 2.9: «departments» class has id, name, created_at, updated_at; «users» class has id, name, email, password, remember_token, created_at, updated_at; «students» has id, character_type, user_id, created_at, updated_at; «teachers» has id, name, user_id, department_id, created_at, updated_at; «subjects» has id, name, department_id, created_at, updated_at; «teacher_subjects» has id, teacher_id, subject_id, created_at, updated_at; «quizzes» has id, title, teacher_id, subject_id, isPsychological, created_at, updated_at; «question_types» has id, name, created_at, updated_at; «student_quizzes» has id, student_id, quiz_id, accepted, result, created_at, updated_at; «questions» has id, title, question_value, quiz_id, question_type_id, created_at, updated_at; «themes» has id, name, subject_id, created_at, updated_at; «materials» has id, title, created_at, updated_at; «answers» has id, right, content, question_id, created_at, updated_at; «question_themes» has id, question_id, theme_id, created_at, updated_at; «theme_materials» has id, material_id, theme_id, character_type, created_at, updated_at and «student_quiz_results» has id, question_id, answer_id.

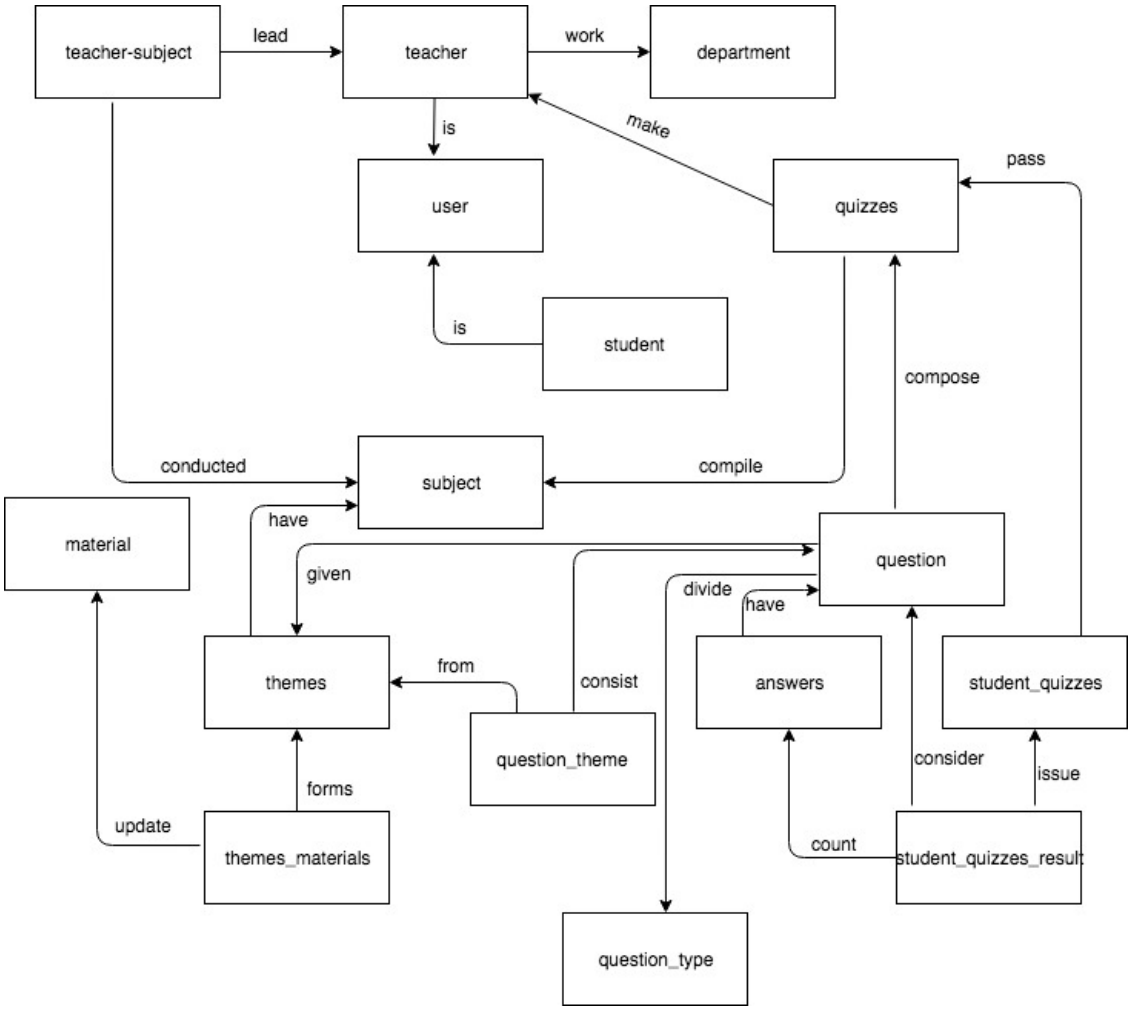

Figure 2.9 – Classes and relationship between them

Each class has subclass or superclass, which defining in Table 2.3. For example, users can be login as student or teacher. In accordance with Figure 2.10 shown OWL classes, especially teachers class.

Table 2.3. Definition of classes for system

Class	***Defining of class***	***Subclass***	***Superclass***
Users	User of system	Student/Teacher	Teacher / student
Departm ent	Department	Department_Name administrator/ Information System	Teacher
Teacher	Teacher	UserName, department Name, Teacher Name	Quizzes, teacher-subject
Quizzes	Psychological test or definition of knowledge	Subject, Teacher, psychological, quizTitle,	Questions, student-quiz, psychological test

Continuation of the table 2.3

Questions	Questions in quizzes	Question_types, quizzes	Answers, question_themes, student_quiz_result
Answers	Answer to the question	Questions, right/wrong, answer_content, question_title, answerID	student_quiz_result
Materials	content	Material_title,material_name	Theme_materials
Subjects	subjects	SubjectName	Themes, quizzes, teacher_subjects

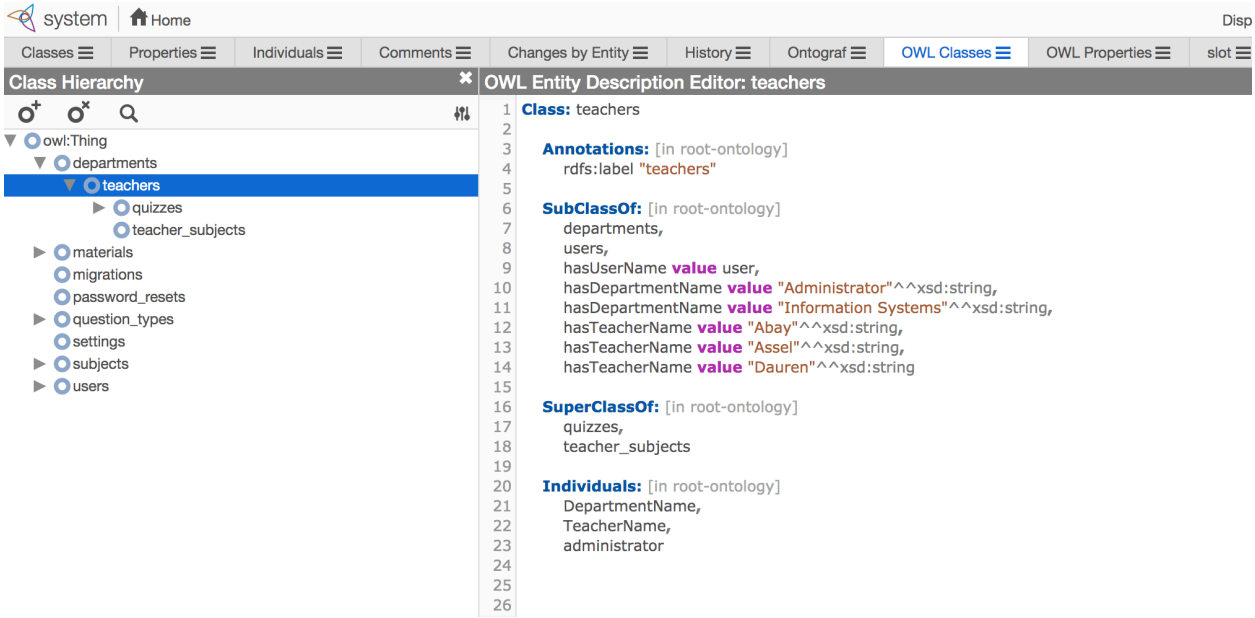

Figure 2.10 – OWL classes

The definition of slots. Slot properties attached to classes. Slots can have facets that describe the type of value, allowed values, number of values (power), and other properties of values that the slot can accept.

- String – the simplest type of value that is used as a simple string;
- Number of float/integer;
- Slot a Boolean – true/false;
- Numbered slots-specifies a list of specific allowed values;
- Instance slots-allow you to define relationships between individual concepts.

In General, several types of object properties can become slots in ontology: internal properties, external properties, parts (if the object has a structure), and relations with other individual concepts. Slots can have different facets that describe the type of value, allowed values, number of values (power), and other properties of values that a slot can accept. Allowed classes for slots of type Instance are often called a range of values of the slot. Classes to which a slot is attached or classes which property a slot describes, are called the domain of the slot. When defining a

domain or range of slot values, find the most common classes or class that can be a domain or range of slot values, respectively. On the other hand, don't define a domain and value range that is too General: all classes in the slot domain must be described by the slot, and instances of all classes in the slot value range must be potential slot placeholders [60]. In Protege owner can change data properties, IRI annotations, characteristics, domain and range in accordance with Figure 2.11.

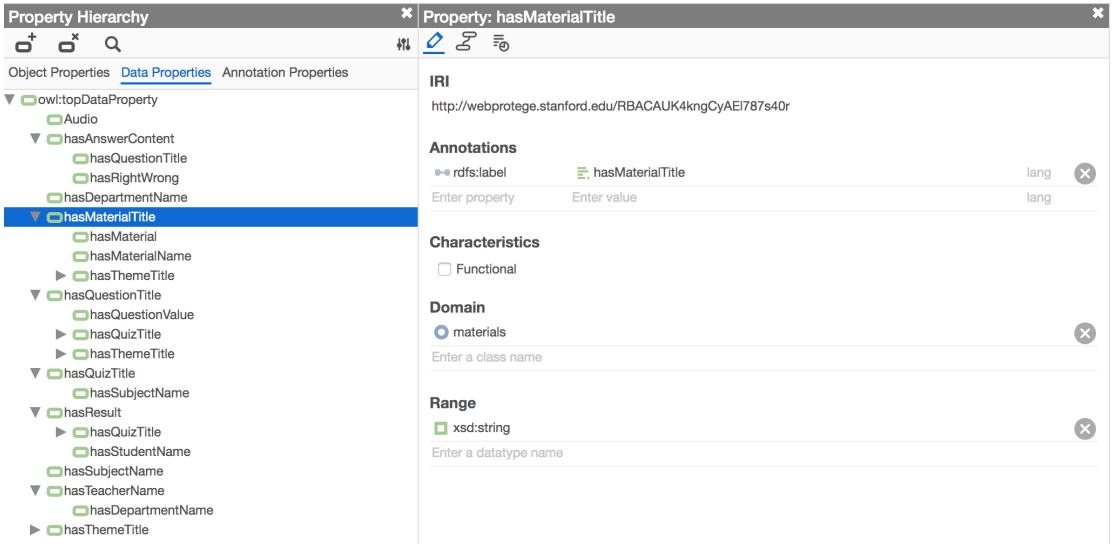

Figure 2.11 – Property hierarchy of slots

Individuals. The last step is to create separate individuals of classes in the hierarchy. To define a separate individuals of a class, owner must do the following actions in accordance with Figure 2.12:

1. Select a class;
2. Create a separate individuals of this class;
3. Enter the values for the slots.

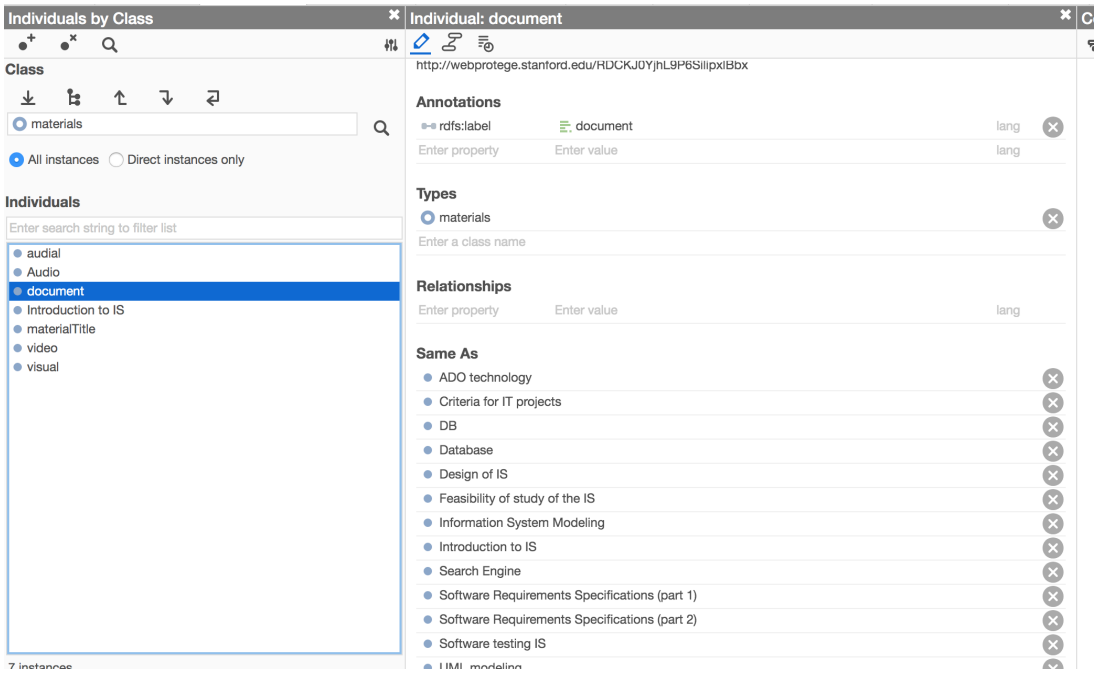

Figure 2.12 – Individuals by Class

Classification of objects based on all the axioms of the ontology, using reasoner (logical inference machine), in other words, we will logically define all the categories that the data belongs to.

Semantic reasoner (semantic reasoning engine) is a piece of software that can deduce logical conclusions (logical consequence) from a set of adequately formalized basic knowledge or axioms.

Query – Manchester, syntax – Protégé.

The Manchester syntax for OWL 2 ontologies is defined using a standard BNF notation, which is summarized in Table 2.4.

Table 2.4 – OWL 2 ontologies

Construct	Syntax	Example
Non-terminal symbols	Boldface	Class Expression
Terminal symbols	Single quoted	'Property Range'
Zero or many	Curly braces	{Class Expression}
Zero or one	Square braces	[Class Expression]
Alternative	Vertical bar	Assertion| Declaration
Grouping	Parentheses	Data property Expression

The syntax for annotations in the Manchester OWL syntax closely corresponds to the syntax in the OWL 2 Functional-Style Syntax. The Manchester syntax is a user-friendly compact syntax for OWL 2 ontologies; It is frame-based, as opposed to the axiom-based other syntaxes for OWL 2 in accordance with Figure 2.13.

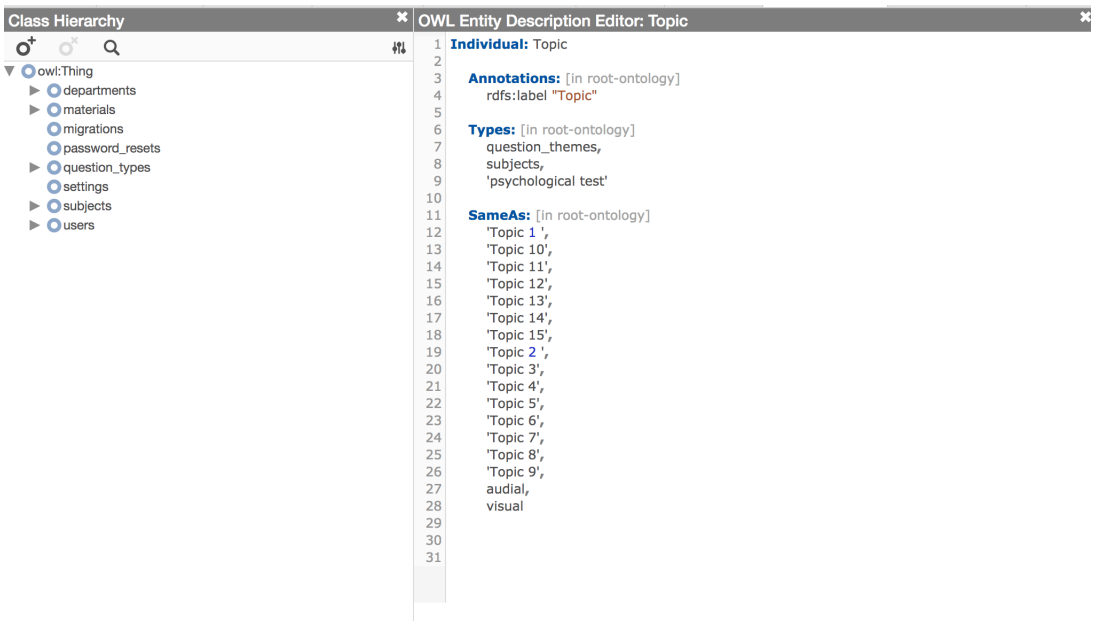

Figure 2.13 – OWL entity

Protégé is a freely distributed Java program designed for building (creating, editing, and viewing) ontologies of a particular application area. It includes an ontology editor that allows you to design ontologies by expanding the hierarchical structure of abstract and concrete classes and slots. Based on the generated ontology, Protégé allows owners to generate knowledge acquisition forms for introducing instances of classes and subclasses. This tool supports the use of the OWL language and allows you to generate HTML documents that display the structure of ontologies. Because it uses the OKBC knowledge representation frame model, it can also be adapted for editing domain models that are not represented in OWL, but in other formats (UML, XML, SHOE, DAML+OIL, RDF/RDFS) [61, 62]. One of the advantages is make a query and see the result of the query in accordance with Figure 2.14.

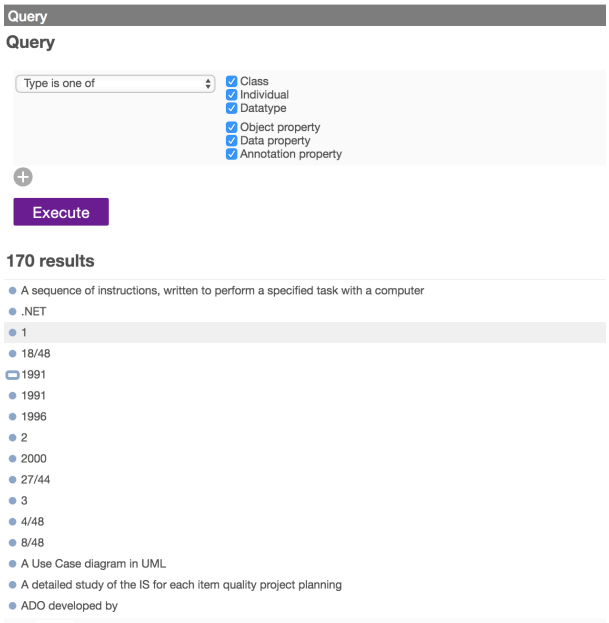

Figure 2.14 – Query

Conclusions on the second section

This section discusses the methodology for creating content using ontology. The development of the ontology begins with the definition of classes, then need to distribute the classes in the taxonomic hierarchy, then the definition of slots and the description of the acceptable values of these slots and at the end filling in the values of the instance slots. Thus, it is possible to create an ontology of content for educational institutions. The distribution of content by the type of user perception is also determined. In education, take into account 2 types of user perception: audial and visual. For the first type of users, more audio materials are needed, and for the second type of perception, video materials are needed. Creating an ontology will help to create conditions for individualizing learning. Using various methods of semantic data analysis, can create an individual learning path for users.

3 ALGORITHMS, MODELS AND METHODS OF INTELLECTUAL INFORMATION AND TRAINING SYSTEM

In this section of the dissertation, various methods and models for calculating the semantic proximity of words and topics are considered. The problem is solved with a sequence of topics in one discipline, as well as with the number of training elements. A hypothesis is proposed, as well as a proof of the hypothesis using various scientific methods. A mathematical model for the information system is developed. The assessment of the effectiveness of training with an intelligent system is also determined.

3.1 Sequence of content generation algorithm

The adaptability of the system is that each user will have their own learning path. Users log on to the system pass 2 tests that will determine the type of perception and level of knowledge for a particular course. The system will provide personalized content each time.

Consider the content generation algorithm. The relationship between the T_N topics of any course can be represented as a directed graph G_T, the vertices of the graph are topics, and the arcs define the relationship between the topics. In accordance with Figure 3.1 illustrate dependence between topics in one course. [63]

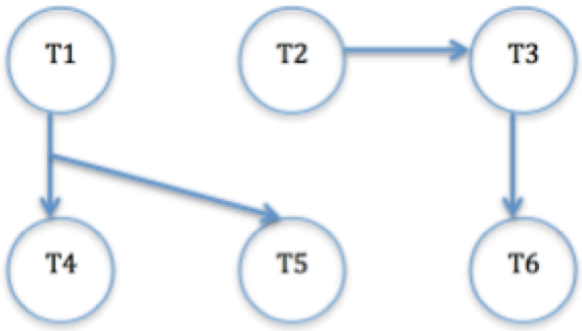

Figure 3.1- Dependence between topics

Properties of the topic graph G_T:

1) There is an initial vertex that corresponds to the beginning of learning process, and an ending vertex that corresponds to the end of learning process.
2) G_T is an infinite graph, since having mastered the sequence of topics, the beginning of which is some topic T_N, the student cannot begin to study it again.
3) Since the graph G_T is infinite, it can be decomposed into levels, that is, represented as a hierarchy. [64]

In this case, at the upper level there will be a fictitious initial vertex, and at the lower level, a fictitious final vertex, which corresponds to the end of the learning process. The final vertex is determined on the basis of the analysis of the competencies that the student must have after graduation.

Consider the algorithm decomposition into levels of an infinite graph G_T [65]:

1) Find a vertex without incoming arcs and assign it a rank r = 0. Delete arcs from this vertex.
2) Suppose that at some stage in the graph there are no vertices without incoming arcs. Assign the following value of rank r to these vertices and cross out the arcs that go out of them.
3) Stage 2 is repeated until all the vertices are ranked.

Remark 1. The infiniteness property is hereditary, that is, if any vertex of the graph is deleted along with incident arcs, the remaining subgraph is also an infinite.

Remark 2. Non-contour graphs have the following properties:

- There is at least one vertex without incoming arcs (initial vertex);
- There is at least one vertex without outgoing arcs (called finite);
- The graph can be represented as a hierarchy, that is, decomposed into levels, while the level number is the length of the maximum path from the initial vertex to the vertex of this level.

Consider the operation of the algorithm for decomposing an infinite graph into levels. In accordance with Figure 3.2 represented by an infinite graph G_T .

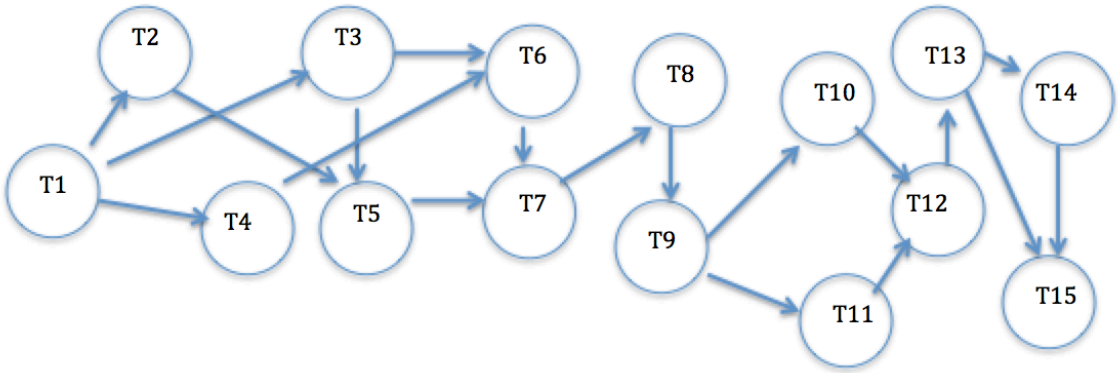

Figure 3.2 - Source infinite graph G_T

Find vertices without incoming arcs and assign them the rank $r = 0$. In this example, the vertices without incoming arcs are the vertices of T_1, in accordance with Figure 3.3. Remove the arcs from this vertex.

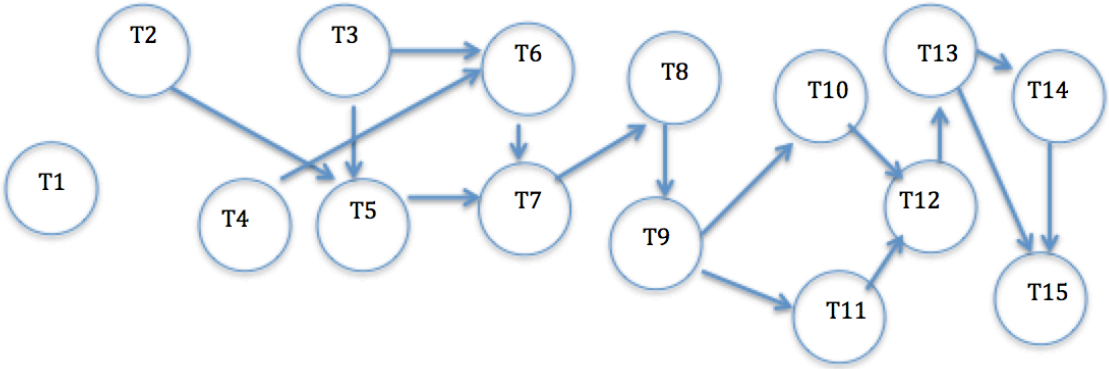

Figure 3.3 – Results of 2nd stage

Find vertices without incoming arcs and assign them a rank $r = 1$. Such vertices are T_2, T_3 and T_4. Remove arcs from vertices. In accordance with Figure 3.4 shows 3rd stage.

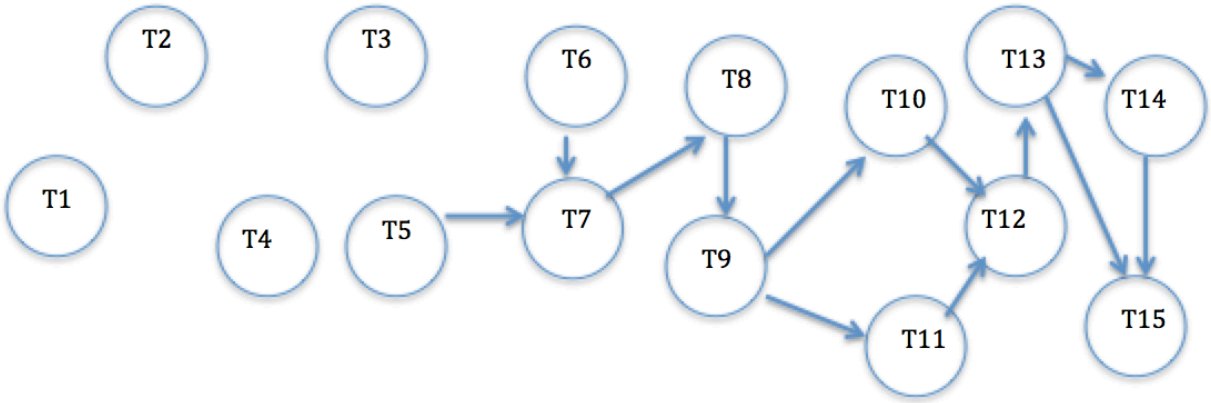

Figure 3.4 – Results of 3rd stage

Find vertices without incoming arcs and assign them a rank r =2. Such vertices are T_5 and T_6. Remove arcs from vertices. In accordance with Figure 3.5 shown results of 4th stage.

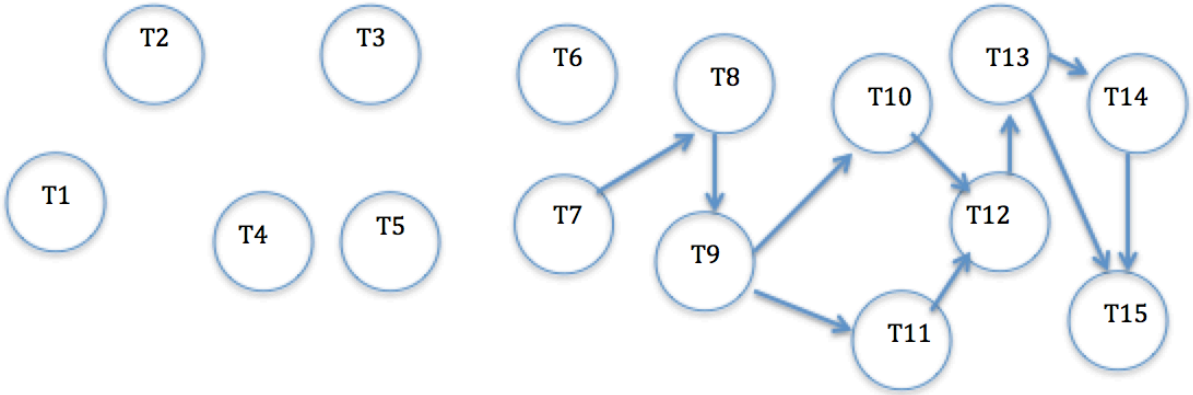

Figure 3.5 – Results of 4th stage

Find vertices without incoming arcs and assign them a rank r =3. Such vertex is T_7. Remove arcs from vertex. In accordance with Figure 3.6 shown results of 5th stage.

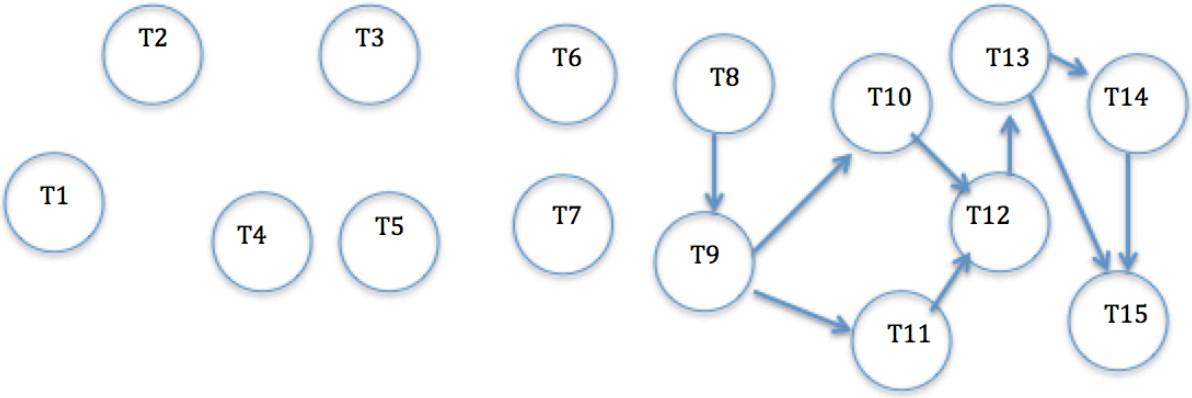

Figure 3.6 – Results of 5th stage

Find vertices without incoming arcs and assign them a rank *r =4*. Such vertex is T_8. Remove arcs from vertex. In accordance with Figure 3.7 represents results of 6th stage.

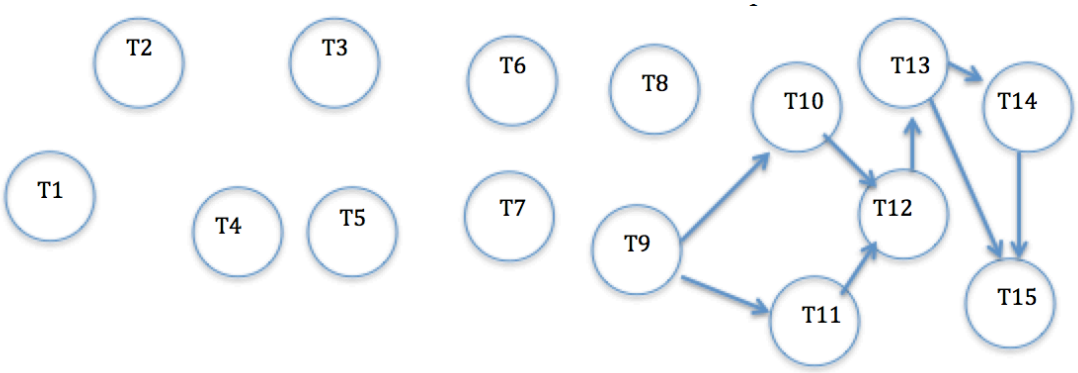

Figure 3.7 – Results of 6th stage

Find vertices without incoming arcs and assign them a rank *r =5*. Such vertex is T_9. Remove arcs from vertex. In accordance with Figure 3.8 represents results of 7th stage.

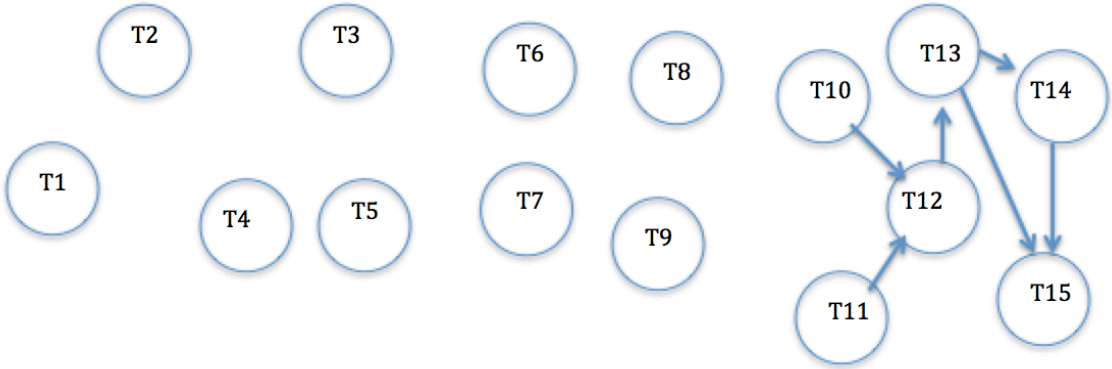

Figure 3.8 – Results of 7th stage

Find vertices without incoming arcs and assign them a rank $r = 6$. Such vertices are T_{10} *and* T_{11}. Remove arcs from vertices. In accordance with Figure 3.9 shows results of 8th stage.

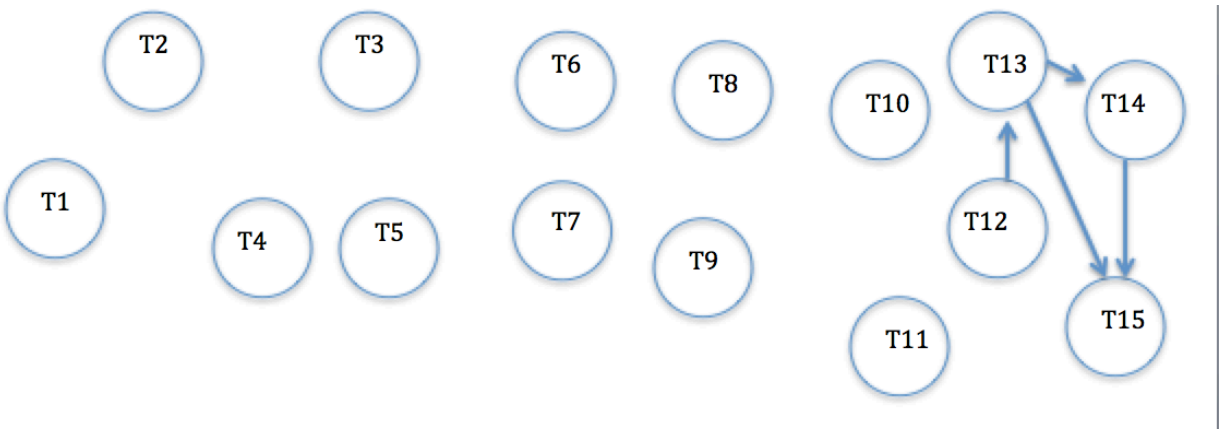

Figure 3.9 – Results of 8th stage

Find vertices without incoming arcs and assign them a rank $r = 7$. Such vertex is T_{12}. Remove arcs from vertex. In accordance with Figure 3.10 shows results of 9th stage.

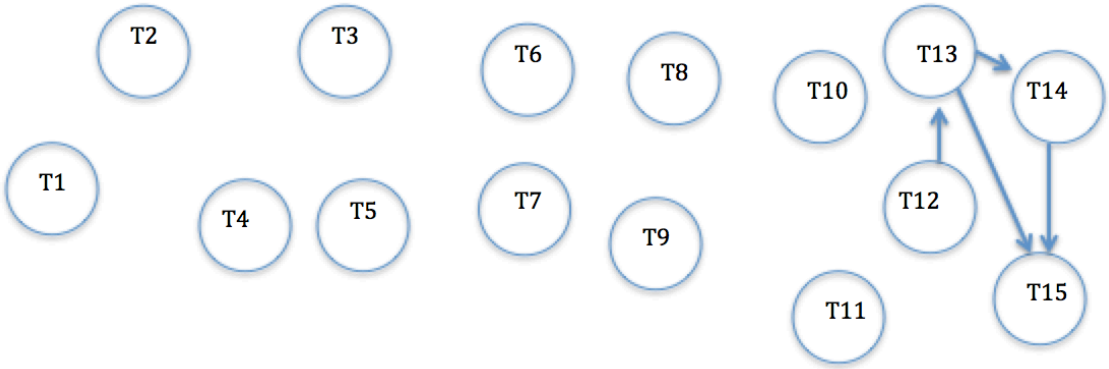

Figure 3.10 – Results of 9th stage

Find vertices without incoming arcs and assign them a rank $r = 8$. Such vertex is T_{13}. Remove arcs from vertex. In accordance with Figure 3.11 represent results of 10th stage.

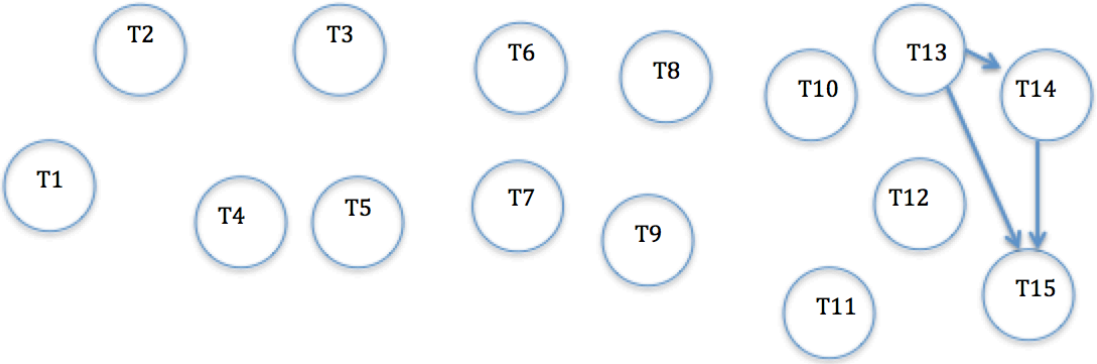

Figure 3.11 – Results of 10th stage

Find vertices without incoming arcs and assign them a rank *r =9* и *r =10*. Such vertices are T_{14} and T_{15}. Remove arcs from vertices. In accordance with Figure 3.12 represent results of 11th stage. Graph is presented hierarchically, in accordance with Figure 3.13.

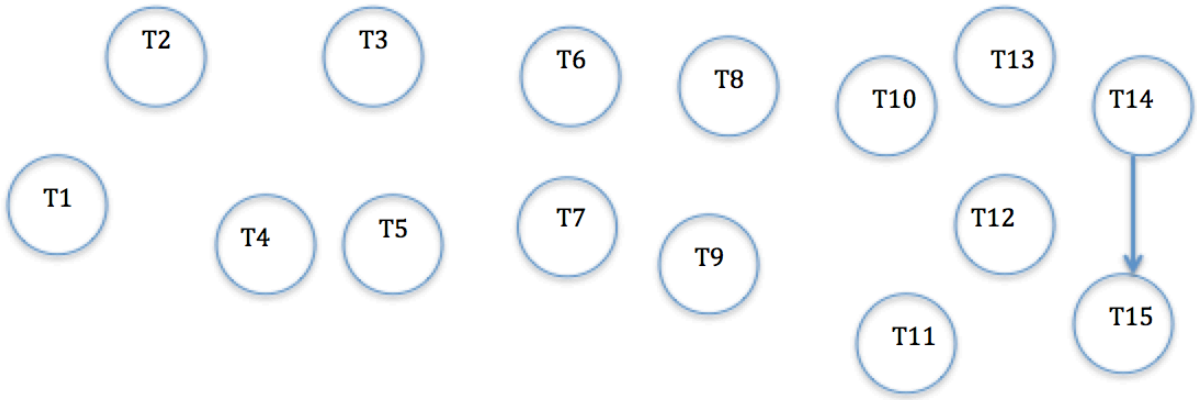

Figure 3.12 – Results of 11th stage

It is planned to obtain an algorithm for the formation of content for an educational system for distance learning, which is adapted for each student, taking into account its features. On the basis of the developed algorithm, an adaptive online educational process is being developed, according to which the forms of education for people with disabilities and persons with disabilities will be developed and alternative forms of content presentation will be created for them. The development and implementation of an online process will make it possible to abandon expensive foreign training systems. [24]

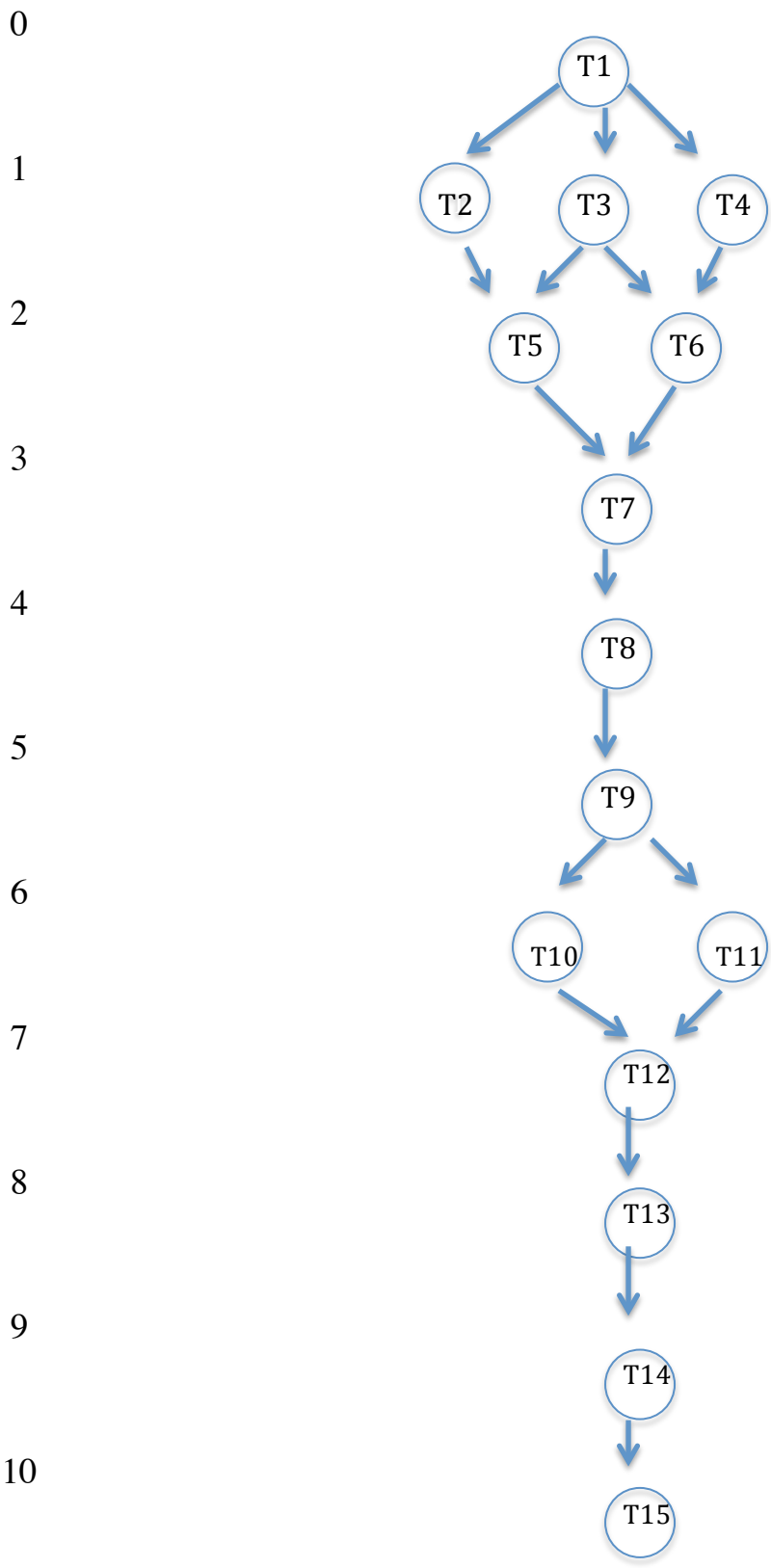

Figure 3.13 – Graph is presented hierarchically

3.2 Determining the semantic proximity of topics using single criteria

Use method for calculating semantic proximity by professor Bondarchuk. Suppose that w_1 and w_2 are words for which semantic proximity must be calculated. The Bondarchuk's method can be divided into several steps:

1. Formation of contextual sets of words w_1 and w_2 (these sets contain words with which the words w_1 and w_2 are often used in the same context) (1-3).

$$c_1\{c_{11}, c_{12}, \dots, c_{1n}\} \quad (1)$$

$$c_2 = \{c_{21}, c_{22}, \dots, c_{2m}\} \quad (2)$$

$$C = c_1 \cup c_2 \quad (3)$$

where C - general contextual set of words.

These sets contain words with which the words w_1 and w_2 are often used in the same context. Then we form a common contextual set of words. Obviously, the cardinality of this set will be equal to n+m.

2. Calculating normalized affinities between the common determinant and each of the words $\boldsymbol{w_1}$ and $\boldsymbol{w_2}$ (4-5).

$$B(c_i, w_1) = \varphi(c_i, w_1)/\max(\varphi(w_1)) \quad (4)$$

$$B(c_i, w_2) = \varphi(c_i, w_2)/\max(\varphi(w_2)) \quad (5)$$

where $\boldsymbol{\varphi(c_i, w_1)}$– number of documents $\boldsymbol{c_i}$ and $\boldsymbol{w_1}$ meet together.

Max.the frequency (w_j) is calculated by the formula as the maximum frequency for all words from the combined context set with (6):

$$\max\left(\varphi\left(w_j\right)\right) = \max\{\varphi(c_i, w_j)\}, c_i \in C \quad (6)$$

where B – proximity, φ – frequency.

3 The calculation of semantic proximity.

Consider the calculation of semantic proximity between the words w_1 and w_2. To do this, we calculate the coefficients R_i for all words from the context set C using the formula (7):

$$R_i = \min\{B(c_i, w_1), B(c_i, w_2)\} / \max\{B(c_i, w_1), B(c_i, w_2)\} \quad (7)$$

where $\boldsymbol{R_i}$ – coefficient of all words from the context set. Then the semantic proximity of the words w_1 and w_2 is calculated by the formula (8):

$$sem(B) = \sum_{i=1}^{k}(p_i * \frac{R_i}{1+R_i} + s)/(1+s) \quad (8)$$

where p_i is the coefficient of joint occurrence in the entire sample, equal to 2 if both words occur in 1 document, otherwise 1. s is the synonymy coefficient equal to 1 if the words are synonyms, otherwise 0. For semantically close words, the coefficients must be close to 1 [66].

As a result of the applied formula, you will get a number in the range [0, 0.75*(n+m)]. To get semantic proximity in the range [0, 1], you need to divide the resulting result by 0.75*(n+m). For semantically close words, the coefficient is close to 1.

Let's calculate the semantic proximity of the two topics on the subject "Fundamentals of information systems", shown in Table 3.1 and Table 3.2.

w_1 = 'Information System modeling';

w_2 = 'UML modeling';

Normalized affinities between the general context set and the words 'IS modeling' and 'UML modeling', w_1-topic of the 4th week, w_2-topic of the 5th week, R_i – coefficient of all words from the context set.

C = {adaptability, integrity, quality, design}

$\boldsymbol{B(C, w_1) = \varphi(C, w_1)/\max(\varphi(w_1))} = 0{,}83$
$\boldsymbol{B(C, w_2) = \varphi(C, w_2)/\max(\varphi(w_2))} = 0{,}25$
$\boldsymbol{B(C, w_1) = \varphi(C, w_1)/\max(\varphi(w_1))} = 0{,}6$
$\boldsymbol{B(C, w_2) = \varphi(C, w_2)/\max(\varphi(w_2))} = 0{,}12$
$\boldsymbol{B(C, w_1) = \varphi(C, w_1)/\max(\varphi(w_1))} = 0{,}32$
$\boldsymbol{B(C, w_2) = \varphi(C, w_2)/\max(\varphi(w_2))} = 0{,}07$
$\boldsymbol{B(C, w_1) = \varphi(C, w_1)/\max(\varphi(w_1))} = 1$
$\boldsymbol{B(C, w_2) = \varphi(C, w_2)/\max(\varphi(w_2))} = 0{,}81$

Table 3.1. - Semantic proximity

words	Information System modeling	UML modeling
Adaptability	0,83	0,25
Integrity	0,6	0,12
Quality	0,32	0,07
Design	1.0	0,81

$$\boldsymbol{R_i = \min\{B(c_i, w_1), B(c_i, w_2)\} / \max\{B(c_i, w_1), B(c_i, w_2)\}}$$

$R_1 = 0{,}31$

$R_2 = 0{,}16$

$R_3 = 0,23$
$R_4 = 0,8$

Table 3.2 – words and coefficient R_i

C	R_i
Adaptability	0,31
Integrity	0,16
Quality	0,23
Design	0,8

s is the synonymy coefficient:
s = 1, if the words are synonymous;
s = 0, otherwise.
$\boldsymbol{p_i}$ is the occurrence coefficient of w1 and w2:
$\boldsymbol{p_i}$ = 2 if both words are in 1 document,
$\boldsymbol{p_i}$ = 1, otherwise.
In our case, s = 0, "Information System modeling" and "UML modeling" not synonyms.

$$sem(B) = \sum_{i=1}^{k}(p_i * \frac{R_i}{1+R_i} + s)/(1+s) \quad = 1,45$$

the range should be [0, 1], if greater then it should be divided by 0.75(n+m).
Sem.proximity = 0.483.

Thus, the semantic similarity between the topics " Information System modeling "and" UML modeling " is 0.483.

Also need to calculate the semantic proximity of all topics with each other, shown in Table 3.3:

Topic 1 – Introduction to the Information Systems
Topic 2 – Software Requirement Specification (part 1)
Topic 3 - Software Requirement Specification (part 2)
Topic 4 – Information System modeling
Topic 5 - UML modeling
Topic 6 - Design of IS
Topic 7 - Database
Topic 8 – Search Engine
Topic 9 - Software Testing Information System
Topic 10 – Feasibility study of the IS
Topic 11 – Criteria for IT projects
Topic 12 – Development of IS

Topic 13 – ADO technology
Topic 14 – Final presentation
Topic 15 – Examination

Table 3.3. Semantic proximity between topics

Topics	w_1	w_2	**Semantic proximity**
Topic 1 - Introduction to the Information Systems	Topic 1	Topic 2	0,3
Topic 2 - Software Requirement Specification (part 1)	Topic 2	Topic 3	0,8
Topic 3 - Software Requirement Specification (part 2)	Topic 3	Topic 4	0,77
Topic 4 - Information System modeling	Topic 4	Topic 5	0,483
Topic 5 - UML modeling	Topic 5	Topic 6	0,81
Topic 6 - Design of IS	Topic 6	Topic 7	0,44
Topic 7 - Database	Topic 7	Topic 8	0,21
Topic 8 - Search Engine	Topic 8	Topic 9	0,37
Topic 9 - Software Testing Information System	Topic 9	Topic 10	0,48
Topic 10 - Feasibility study of the IS	Topic 10	Topic 11	0,91
Topic 11 - Criteria for IT projects	Topic 11	Topic 12	0,27
Topic 12 - Development of IS	Topic 12	Topic 13	0,67
Topic 13 - ADO technology	Topic 13	Topic 14	0,34
Topic 14 - Final presentation, Topic 15 - Examination	Topic 14	Topic 15	0,94

Put-forward hypothesis: if the semantic proximity is greater than 0.75 (as a result of expert surveys and an experiment), then the topics are dependent, namely, a large semantic relationship between the topics is determined. In other words, the content of these topics is linked and dependent on each other. And of course, the themes are consistent. If we stick to the hypothesis, then the dependence appears in the themes:

1. dependency - T2, T3;
2. dependency – T3, T4;
3. dependency – T5, T6;
4. dependency – T10, T11;

5. dependency – T14, T15.

Rules:

- Rule 1: If the student has not mastered the 2nd topic, they will not be able to go to the 3rd topic;
- Rule 2: If a student has mastered topic 2 but has not mastered topic 3, they will not be able to go to topic 4;
- Rule 3: If a student has not mastered topic 5, they will not be able to go to topic 6;
- Rule 4: If a student has not mastered topic 10, they will not be able to go to topic 11;
- Rule 5: If a student has not mastered topic 14, they will not be able to go to topic 15.

Using semantic representation of data, as well as semantic analysis of the text, can develop a sequence of topics in a single subject. And in the future, can use this method not only in the subject, but also in the course, or in drawing up a plan for academic years.

3.3 Determining the proximity of topics by many criteria using a scientific experiment

The experiment was conducted at Astana IT University during the first trimester of 2019 year among first-year students. This experiment involved 115 users divided into 5 groups. Students first took a test to determine the level of knowledge, then a test to determine the type of perception. There are only 2 types of perception for the educational system: visual and audial. The type of perception determined the type of content. For visuals, the system provided more videos, presentations, for audials, only audiobooks and lectures. After each lecture, tasks were given to complete, as shown in Table 3.4. During for the first trimester users can pass all asignments, include practical or laboratory tasks and quizzes. Quizzes was multiple questions choice. All asignments given by order.

Table 3.4 – all points for one group

		type	level kn	a1	a2	a3	a4	a5	a6	a7	a8	a9	a10	a11	a12	a13	a14	a15
1	1_student	audial	22	95	49	64	70	72	78	85	93	87,5	95	90	80	81	87	82,5
2	2_student	visual	20	95	63	73	55	74	76	85	85	75	85	80	70	74	75	75
3	3_student	visual	20	95	43	50	60	57	72	85	92	87,5	100	75	80	80	90	81,6
4	4_student	audial	23	70	43	53	65	32	53	85	85	37,5	75	70	50	62	74	64,1
5	5_student	audial	22	95	60	60	70	88	83	80	90	56	85	85	70	73	84	80,4
6	6_student	audial	16	95	76	73	69	75	85	90	90	75	75	77	60	75	100	88

Continuation of the table 3.4

7	7_student	visual	15	55	76	70	59	30	44	60	74	62,5	85	80	60	64	50	52,4
8	8_student	audial	15	95	76	65	50	90	84	70	85	87,5	100	90	70	75	100	87,7
9	9_student	audial	15	95	43	33	75	63	73	50	50	62,5	80	85	60	60	65	65,9
10	10_student	audial	7	0	66	70	83	71	75	80	75	0	85	80	60	63	77	72,2
11	11_student	audial	13	80	76	70	59	57	61	75	70	50	70	75	70	69	64	64,6
12	12_student	visual	17	85	76	64	60	85	85	90	90	87,5	90	85	85	83	85	84,4
13	13_student	audial	13	0	43	53	49	52	54	75	80	0	0	75	70	59	60	57,9
14	14_student	audial	21	95	46	57	83	78	77	60	90	0	100	75	55	90	99	89,7
15	15_student	visual	29	95	76	70	60	73	77	85	85	75	85	90	80	79	85	80,8
16	16_student	audial	17	85	60	53	70	73	78	85	90	87,5	90	90	80	79	79	78,7
17	17_student	visual	18	80	63	67	60	85	75	80	90	0	0	60	70	90	90	85,5
18	18_student	audial	28	85	49	50	66	93	83	70	84	87,5	85	85	85	78	80	80,3
19	19_student	audial	7	95	49	57	64	56	77	75	75	75	85	90	75	75	79	77,2
	Average			78,42	59,63	60,63	64,6	68,63	73,15	77,1	82,78	60,75	81,66	80,89	70	74,15	80,15	76,25

Proof of the hypothesis. H_o - null hypothesis, all topics are unrelated. H_A - alternative hypothesis, topics are related, where the semantic proximity is greater than 0.75. Using local reference methods such as the Dice measure, the Jaccard index, and the Kulchinsky measure, and the Pearson correlation method, one can characterize the existence of a linear relationship between two quantities.

Pearson's correlation method. Despite its widespread application, Pearson's correlation method has many limitations raising the question of whether its use should be restricted or even avoided. First, correlation method tests only whether there is a linear relationship between two variables and a significant curvilinear relationship can result in a non−significant correlation method. Second, the use of correlation method assumes that the pairs of observations (x,y) are members of the bivariate normal distribution and failure of this assumption may require the use of a non−parametric correlation coefficient. Third, the square of the correlation coefficient, also known as the 'coefficient of determination', measures the proportion of the variance associated with the Y variable that can be accounted for or 'explained' by the X variable [67]. The following equation (1) presents Pearson's correlation method:

$$K_{xy} = \frac{\sum\left(x_i - \bar{x}\right)\cdot\left(y_i - \bar{y}\right)}{\sqrt{\sum\left(x_i - \bar{x}\right)^2 \cdot \left(y_i - \bar{y}\right)^2}} \quad (1)$$

where x_i - value accepted by the X variable;
y_i - value accepted by the Y variable;
$\bar{x}$ - average value for the X variable;

$\bar{y}$ - average value for the Y variable;

Dice's similarity measure. Dice's coefficient simply measures the words that two texts have in common as a proportion of all the words in both texts, shown in equation (2) [68]. Let A, B be the concepts under consideration, and n(A) and n (B) be the numbers of vertices incident to A and B, respectively.

$$sim(A,B) = \frac{2|n(A) \cap n(B)|}{|n(B)| + |n(B)|} \quad (2)$$

Jaccard's Index. The Jaccard's index is a well-known measurement of the similarity between two sets A and B, defined as the size of the intersection divided by the size of the union of the two sets, presents in equation (3):

$$sim(A,B) = \frac{|n(A) \cap n(B)|}{|n(A) \cup n(B)|} \quad (3)$$

where we say sim(A, B)=1 if |A∪B|=0. By reinterpreting a set of patterns as a set of elements, we are able to use the Jaccard index to measure the similarity between two sets of patterns [69].

Kulchinsky's measure. Local reference measures include Kulchinsky's measure, as equation (4) [70]:

$$sim(A,B) = \frac{2|n(A) \cap n(B)|}{|n(B)| + |n(B)| - 2|n(A) \cap n(B)|} \quad (4)$$

Comparison between methods. Teachers gave an expert assessment of the relationship between the topics of each lecture. We need to calculate of semantic proximity using the Bondarchuk's method, the Jaccard's index, the Kulchinsky's measure, the Dice's measure, expert evaluation and the Pearson's correlation. Students passed the tasks and average values were used for all methods. A summary of the full text was taken from each topic. By the number of words between topics, we can calculate the proximity of words using different methods. A summary of the first topic: "A computer program is a sequence of instructions written to perform a specified task with the help of a computer", summary of the second topic: "Software is a set of computer programs, procedures and related documentation and data (ISO/IEC 12207)" and summary of the third topic: "The basic organization of the system embodied in its hardware, software and information components and their relationships with each other and with the environment, as well as the principles governing the IS design and development". Summary of the fourth topic: "The ability

of Information System to make this or that decision based on the knowledge knowledge of the user collected in the course of their work on PC", summary of the fifth topic: "Model always includes only essential elements which affect the results, and doesn't include non-essential elements, that do not affect the results" and summary of the sixth topic: "To sum up, this lecture focuses on two main types of models: structural or static and behavior or dynamic". Summary of the seventh topic: "A RDB is a DB /organized in accordance with certain rules/ and maintained in the computer's memory /as a set of logically related data/ that characterizes the current state of a subject area/ and is used to meet/ the information needs of users", for eighth topic: "Search engines are programs that search documents for specified keywords and returns a list of the documents where the keywords were found" and for ninths topic: "Study, test, and analysis of information system gives an opportunity to identify errors in the software and evaluate its quality. Based on this knowledge, you can debug (correct, improve) Information System, which will make its competitive ability, cost-effective and efficient". Summary of the tenth topic: "A feasibility study of the project (feasibility study) is necessary in the event of a major investment for the development of business IT", for eleventh topic: "Global changes at informational, communicational, professional and other spheres of modern society require correction of informative, methodical, technological aspects of education, review of former value priorities, targets and pedagogical assets" and for twelfths topic: "The meaning of education consists in development of students' independent problem solving ability at different spheres and kinds of activity on the base of social experience usage, which includes students' own experience". Summary of the thirteenth topic: "Establish communication between the conclusion and the beginning of the performance. At the end of your speech emotion should not be lower emotionality your entry", for fourteenth topic: "Aims to mastering of physical, spiritual, and intellectual self-development approaches, emotional self-regulation and self-support by student" and last one, summary of the fifteenth topic: "Object – information source for researcher, that objectively exists in theory and practice. Subject more concrete. It is subjected to direct analysis in research". Next step, count the number of words in each summary:

- 1st topic – 18;
- 2nd topic – 15;
- 3rd topic – 35;
- 4th topic – 28;
- 5th topic – 22;
- 6th topic – 19;
- 7th topic – 38;
- 8th topic – 23;
- 9th topic – 36;
- 10th topic – 23;
- 11th topic – 30;
- 12th topic – 30;

- 13th topic – 25;
- 14th topic – 19;
- 15th topic – 23.

To determine the proximity of topics, need to count the number of identical words in each summary, shows in table 3.5.

Table 3.5 - Determination of proximity of topics

Topics	**Number of words for 1 topic**	**Number of words for 2 topic**	**Identical words between 2 topics**
T1 – T2	18	15	4
T2 – T3	15	35	3
T3 – T4	35	28	5
T4 – T5	28	22	2
T5 – T6	22	19	1
T6 – T7	19	38	2
T7 – T8	38	23	2
T8 – T9	23	36	1
T9 – T10	36	23	1
T10 – T11	23	30	1
T11 – T12	30	30	4
T12 – T13	30	25	1
T13 – T14	25	19	1
T14 – T15	19	23	1

For expert evaluation 3 teachers of the same discipline evaluated the proximity of topics. Each topic was evaluated and the final score was taken as an average value.

Substituting for the method formulas we get the values for all methods. The Table 3.6 shows pairs of topics and also calculated semantic proximity using the Bondarchuk's method, the Jaccard's index, the Kulchinsky's measure, the Dice's measure, expert evaluation and the Pearson's correlation.

Table 3.6 - Semantic proximity of topics using different methods

w_1	w_2	**Bondarchuk's semantic proximity**	**Jaccard' index**	**Kulchinsky' measure**	**Dice' measure**	**Expert evaluation**	**Pearson's correlation**
Topic 1	Topic 2	0,3	0,1379	0,2244	0,2666	0,3	0,05246
Topic 2	Topic 3	0,8	0,0638	0,1428	0,0857	0,9666	0,77802
Topic 3	Topic 4	0,77	0,0862	0,1607	0,1785	0,75	0,7516
Topic 4	Topic 5	0,483	0,0416	0,0881	0,0909	0,4666	0,06282

Continuation of the table 3.6

Topic 5	Topic 6	0,81	0,025	0,049	0,0526	0,9	0,89231
Topic 6	Topic 7	0,44	0,0363	0,0789	0,0526	0,45	0,28259
Topic 7	Topic 8	0,21	0,0338	0,0697	0,0869	0,3166	0,68234
Topic 8	Topic 9	0,37	0,0172	0,0356	0,0277	0,3166	0,12597
Topic 9	Topic 10	0,48	0,0172	0,0356	0,0434	0,55	0,55783
Topic 10	Topic 11	0,91	0,0192	0,0384	0,0333	0,883	0,82699
Topic 11	Topic 12	0,27	0,0174	0,1333	0,0133	0,25	0,46811
Topic 12	Topic 13	0,67	0.0185	0,0366	0,04	0,5333	0,44746
Topic 13	Topic 14	0,34	0,0232	0,0463	0,0526	0,3666	0,73213
Topic 14	Topic 15	0,94	0,0243	0,048	0,0434	0,9166	0,95878

The correlation and semantic proximity show that they are closer to the expert assessment of teachers. Thus, we can conclude that where there is a semantic proximity greater than 0.75, the topics are dependent and consistent. For this subject, there are 5 pairs of topics that are interrelated, as described in the hypothesis. In accordance with Figure 3.14 shows how the semantic proximity is greater than 0.75 according to the Bondarchuk's method and the Pearson's correlation there is a similarity of data.

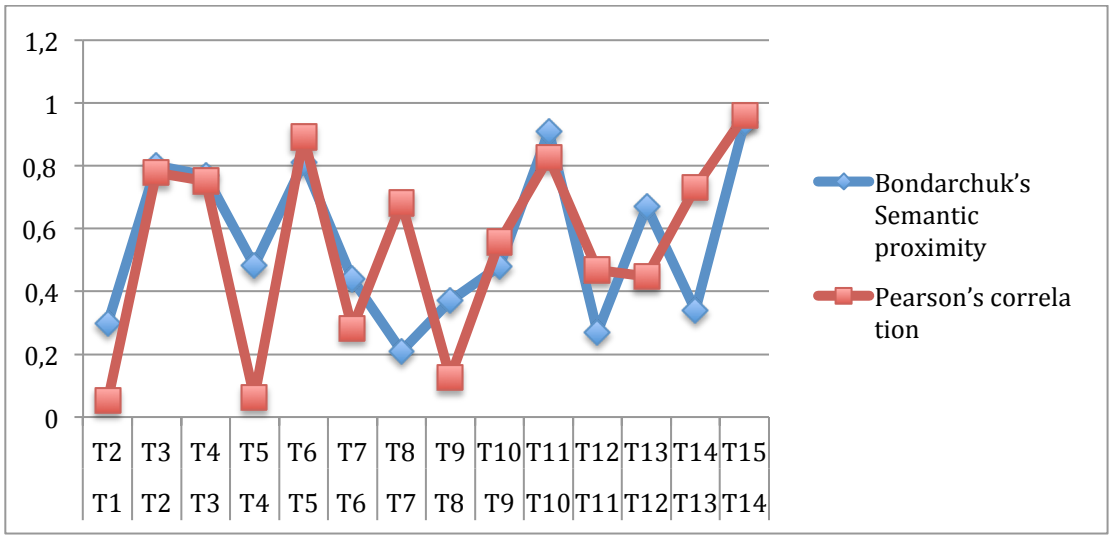

Figure 3.14 - Chart of the relationship between topics and methods: Bondarchuk's semantic proximity and Pearson's correlation

During testing, users used an intelligent system for training. Through the system, we received completely content based on the type of perception. Students also took tests and assignments. Thus, we can see the initial level of knowledge and the final knowledge, in accordance with Figure 3.15. Grades for assignments are also similar to the semantic proximity of topics.

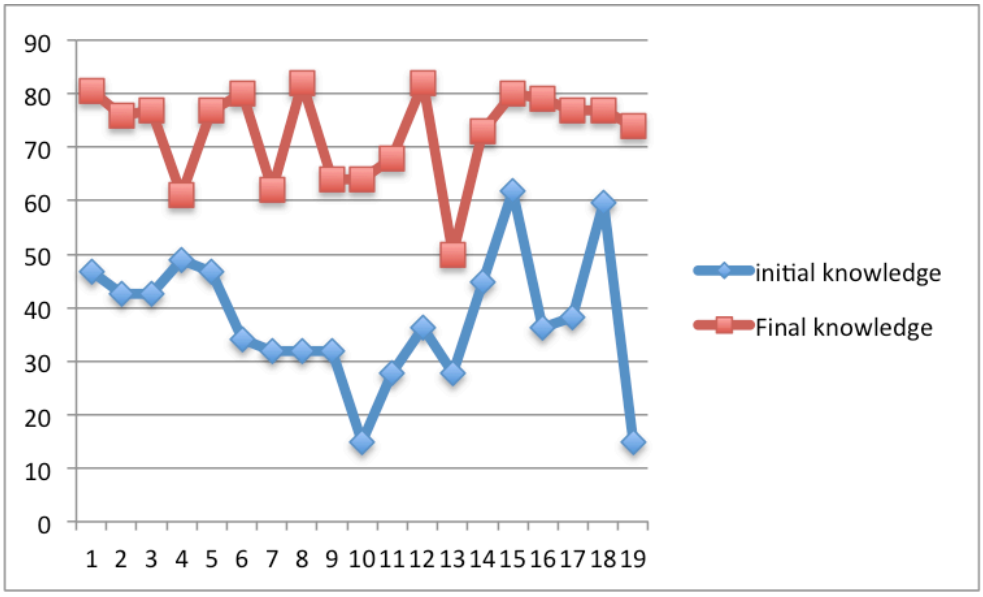

Figure 3.15 - Chart of the relationship between initial knowledge and final knowledge

Semantic analysis, as well as methods of semantic proximity of topics, helps to solve many problems in the educational sphere: the sequence of topics, the number of educational elements and the content. [71]

3.4Mathematical model for information and training system

Mathematical model has been developed for information and training system, which includes 4 different coefficient (5) and has condition (6).

$$W = \frac{1}{4}(\alpha w_1 + \beta w_2 + \gamma w_3 + \delta w_4) \qquad (5)$$

where $\alpha -$ coefficient of Bondarchuk′ semantic proximity,
β – coefficient of Jaccard' index,
γ - coefficient of Kulchinsky' measure,
δ - coefficient of Dice' measure.

The condition characterizes the put-forward hypothesis: if any x(i) word respond in Text T1, y(j) respond in Text T2 and semantic proximity of two topics is equal to or exceeds 0,75, then these two topics are similar in meaning, otherwise the two topics are different.

$$\text{If } \forall(x(i) \in T1)\ and\ ((y(j) \in T2)\ and\ (\rho(x(i), y(j))) \geq 0{,}75$$
$$\Rightarrow\ (x(i), y(j) \subseteq z) = 1$$
$$else\ \ (x(i), y(j) \nsubseteq z) = 0 \qquad (6)$$

where x(i) – i a word in the text T1

y(j) – j a word in the text T2
ρ (·) – distance calculated using the Bondarchuk method
$\rho(\cdot) \geq 0{,}75$

3.5 Extracting knowledge for Microlearning

Microlearning is a skill-based approach to learning that delivers information in small, highly focused chunks. It is the ideal way to find quick answers to specific problems. [72]

Microlearning can help people to expand their knowledge and not waste a lot of time. To create the content of education in microlearning, it is necessary to extract knowledge and use annotations, thus compressing the main facts and obtaining the necessary information.

Annotation is an object that is assigned to a piece of text and describes the properties of this piece.

Model - fact extraction scheme (FES) (7-9):

$$<A, Res, C> \quad (7)$$

where $A = \{c_1, \ldots, c_n\}$

$$c_i = <t, s, h> \quad (8)$$

t – specifies the type of the element,
s - semantic entity class,
h – parameter governing the use of the inheritance hierarchy defined for entity classes.

$$Res = <t, op(t), P> \quad (9)$$

Res - result of application FES,
where op(t) – type of transaction,
P – many rules for forming / editing an object,
C – the set of restrictions imposed on the characteristics of the fact arguments.

Microlearning should have limitations:

- Conditions for the semantic characteristics of schema arguments
- Limiting the syntactic compatibility of syntactic group vertices that implement schema arguments
- Structural-textual restrictions on the mutual arrangement of arguments in the text: position of arguments relative to each other, type of contact, type of segment.
- After extracting the facts, you can proceed to compiling microlearning

Advantages of Microlearning:

1) Concentration of attention, since users receive small content, the materials are absorbed faster;
2) Availability, since the materials are provided in an online format;
3) Mobility, as an adaptation to new knowledge and active use of the material;

4) Modularity and flexibility, since the content consists of blocks, will also be provided depending on the type of user perception, which gives the user the opportunity to master more material.

According to Microsoft research, the average attention span is now just 8 minutes. Microlearning should contain content, especially the duration of video or audio materials should be 5-10 minutes, number of slides in the presentations should be 5-10, number of pages in the text materials should be 1-2 pages, test and practical tasks.

Nowadays Microlearning is relevant and this method is used in all areas, for example: Ed.ted.com, dictionary.com, youtube, duolingo.com, Coursmos, udemy.com etc.

The structure of Microlearning is similar to the structure of the traditional teaching method, but there are differences, such as testing of the initial level of knowledge, lesson (audio/video and text material), test or practical tasks and final test/ practice task. Each user category has its own mini-course. The total number of the category is 10.

Many researchers have studied the use of micro-education in education. Oksana Chulanova, doctor of Economics, there are works on the topic "micro-education as a technology for improving the training of the organization's personnel in order to obtain targeted knowledge". Anna Avramenko, a scientist from Moscow State University, has works on the topic "individualization of the process of forming foreign language lexical competence based on micro-learning". Thus, with the help of micro-training, you can teach users any course in any direction.

3.6 Assessment of the efficiency of using intellectual information system

To implement the performance assessment, 2 groups were created:

1) Experimental group using intellectual information system;
2) Control group - traditional method.

The degree of content assimilation is calculated.

The significance of differences between the control and experimental groups was assessed by a nonparametric method of statistical analysis using the Mann-Whitney test. The criterion is designed to test the validity of the differences between two independent samples by the level of the attribute measured on the order scale.

The general idea of the method is that the values of the attribute are assigned ranks, and the ranking is carried out simultaneously for both samples. Then, the experimental value of the U - test is calculated from the ranks, which reflects the degree of overlap of the intervals of the rank values in the two samples; the smaller the Uexc, the smaller the overlap of the intervals and, therefore, the more likely that the difference is significant, as shown in Table 3.7. To test the hypotheses, the U_{exp} is compared with a tabular critical value (selected depending on the sample size and statistical significance): if $U_{exp} > U_{cr}$, H0 is assumed, otherwise H1 is assumed. Limitations of the applicability of the U-test:

1) the volume of samples must be at least three (n1, n2≥3); it is allowed to have only two observations in one of the samples, but in the second there must be at least five (n1 =2, n2 ≥5);

2) the volume of each sample should not exceed 60 (this is due to the limited critical value tables).

From qualitative or quantitative gradations of a trait, it is often necessary to move on to numbers that characterize the order of gradations-they are called ranks.

$$U_{exp} = n_1 * n_2 + \frac{n_m(n_m+1)}{2} - R_m \quad (10)$$

where n - the number of students in the group,
R – the sum of the ranks in a certain group,
R_m – maximum value of R,
U - critical value of the Mann-Whitney U-test. [73]

Table 3.7 – Calculations of assessment of the efficiency

No	*Group*	*Name*	*Point*	*Rank*	*General number*
1	1	7_student	26	1	1
2	2	2_1_student	28	2	2
3	1	4_student	45	3	3
4	2	2_2_student	50	4	4
5	1	13_student	52	5	5
6	2	2_3_student	55	6,5	6
7	1	2_student	55	6,5	7
8	2	2_4_student	57	8	8
9	2	2_5_student	60	10	9
10	2	2_6_student	60	10	10
11	1	15_student	60	10	11
12	2	2_7_student	63	12,5	12
13	1	8_student	63	12,5	13
14	2	2_8_student	65	16	14
15	2	2_9_student	65	16	15
16	2	2_10_student	65	16	16
17	1	1_student	65	16	17
18	1	17_student	65	16	18
19	2	2_11_student	66	19	19
1	2	2_12_student	68	30,5	20
2	1	16_student	68	30,5	21
3	2	2_13_student	70	23	22
4	2	2_14_student	70	23	23

Continuation of the table 3.7

5	1	3_student	70	23	24
6	2	2_15_student	75	26	25
7	1	5_student	75	26	26
8	1	11_student	75	26	27
9	1	18_student	78	28	28
10	2	2_16_student	84	29,5	29
11	2	2_17_student	84	29,5	30
12	1	14_student	87	31	31
13	1	9_student	89	32,5	32
14	1	19_student	90	34	33
15	1	6_student	91	35	34
16	1	10_student	92	36	35
17	1	12_student	97	37	36

$n_1 = 19$;
$n_2 = 17$;
$R_1 = 409$;
$R_2 = 281,5$;
$R_m = 409$;
$n_m = 19$;
$U_{exp} = 104$;
$U_{cr} = 123$;

In our case, $U_{exp} < U_{cr}$ therefore, it is accepted H_1 - there is a significant difference in the levels of assimilation in the compared groups.
In group 1, they coped with the proposed task faster.

It was found that in the experimental group the degree of assimilation increased by 0.62 points, and in the control group it fell by 0.13 points. The effectiveness of training with the use of IS is 1.15 times higher.

Verification of the reliability of the obtained results showed that in the expert group, the result of the analic re-cut will lie in the area: from 3.53 to 4.15 points, and in the control group: from 3.2 to 3.33.

Conclusions on the third section

This section discusses algorithms, models and methods for creating an intellectual information system for education. The level decomposition algorithm is used to solve the content sequence. A hypothesis is put forward and proofs are given using various methods, including the calculation of the semantic proximity of Bondarchuk's method, Jaccard index, the Kulchinsky measure, the Dice measure, expert evaluation and the Pearson's correlation. A mathematical model of the learning system based on the semantic proximity of topics is proposed. It is used to

extract knowledge for the microlearning method. Microlerning allows users to perceive and assimilate information faster and easier. The evaluation of the efficiency of using the system using a nonparametric method of statistical analysis using the Man-Whitney test is calculated. The effectiveness of training with the use of IS is 1.15 times higher.

4 TECHNICAL IMPLEMENTATION OF AN INTELLECTUAL INFORMATION AND TRAINING SYSTEM

This section describes the technical implementation of an intelligent information and training system “edu-elt.kz”. The software is implemented as a website using the open-source general-purpose programming language PHP, the MySQL integration DBMS, and a web framework, Laravel, designed for development using the MVC architectural model. The choice of the PHP programming language is due to the fact that it is a scripting language created for sites with dynamic content.

4.1 Information system development methodology

The system development methodology consists of the following stages:

1) Identification
2) Testing
3) Conceptualization
4) Learning process

The sequence of steps for building an information system in accordance with Figure 4.1.

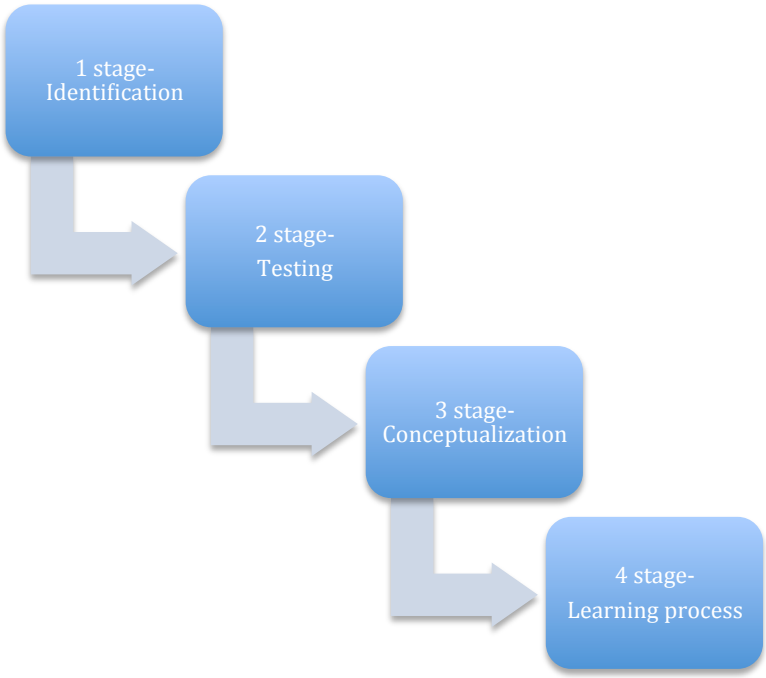

Figure 4.1 - Stages of IS development

4.1.1 Authentication stage

The authentication stage begins with the user's registration. Next comes the authorization and authentication process. Based on the authentication factors, the user

ID is determined. In the information and training system, the identification option is the user identification, and the authentication factor is a username and password. The result of the identification, namely the ID is the user's login.

4.1.2 Testing stage

During this stage, the knowledge is checked and the type of perception of the user is determined. After the identification stage, the system offers courses to pass. When selected, a test is given to determine the level of knowledge for a certain course. For example, when choosing an Educational Practice course " Fundamentals of Information System", the test consists of 28 questions, which includes all 15 topics for this course. The database stores 102 questions, with 10 questions for each topic, including 3 types of question difficulty and 5 answer options. Example of a database of questions:

$$$0030 Aims to mastering of physical, spiritual, and intellectual self-development approaches, emotional self-regulation and self-support by student
A) Value-sense competency
B) Socio-labor competency
C) Informational competencies
D) Actors and use case diagram
E) Personal self-improvement competency
{Correct answer}=E
{ Complexity }=2
{Topic}=12

On this question, you can understand that 1 question is given 5 possible answers, where the correct answer is only 1. Difficulty - 2 and topic-12. Correct answer questions basically, users were given questions on the medium and complex level on each topic in a random order. If the user answered the average and difficult question correctly, then he moved on to the next topic. That is, if the user answered correctly the 1st medium-difficulty question and the 2nd difficult-difficulty question, then the user was asked to start with the topic where he did not answer correctly. Thus, the user is given only the content that he does not know. And the topics that the user knows can also be viewed, but there will be no such topics in the learning path.

4.1.3 Conceptualization stage

At this stage, a meaningful analysis of the problem area is carried out, the concepts used and their relationships are identified. This stage ends with the creation of content for the educational system and the domain model, which includes the main concepts and relationships. At the stage of conceptualization, the following features of the task are determined:

- types of available data;

- source and output data;
- applied hypotheses;
- types of relationships between objects;

The information system for education has 19 entities and their relationships. There are two approaches to the process of building a domain model at the conceptualization stage:

1. The attribute or attribute approach presupposes the presence of information received from experts in the form of triples "object-attribute-attribute value", as well as the presence of training information.
2. The structural or cognitive approach is carried out by highlighting the elements of the subject area, their interrelationships and semantic relations.

4.1.3.1 Attributive approach to building a domain model

A feature-based or attribute-based approach. The attribute approach is characterized by having the most complete information about the subject area: about objects, their attributes, and attribute values. The idea of the method is that rules (combinations of attribute values) are constructed to distinguish one object from another.

4.1.3.2 Structural approach to building a domain model

The structural approach to the construction of the domain model involves the allocation of the following cognitive elements:

1. Concepts;
2. Relationships;
3. Meta-concepts;
4. Semantic relations.

Concepts of the subject area. The concepts of the subject area should form a system that has the following properties:

- uniqueness (lack of redundancy);
- completeness (a fairly complete description of various processes, facts, phenomena of the subject area);
- reliability – validity-the correspondence of the selected units of semantic information to their real names);
- consistency.

Methods of constructing a system of concepts:

- Local Representation method;
- Method for calculating the utilization rate;
- Method of forming a list of concepts;
- Method for making a list of elementary actions;
- Textological method.

Relationships between the concepts of the subject area. A group of methods for establishing relationships involves establishing semantic proximity between

individual concepts. At the heart of establishing relationships is the fundamental category of proximity of objects or concepts.

Meta-concepts of the subject area. Interpretation is usually easier for the expert if the groupings are obtained by informal methods.

The establishment of semantic relations between the concepts of the subject area. The last stage of building a domain model in conceptual analysis is the establishment of semantic relations between the selected concepts and meta-concepts. To establish semantic relations means to determine the specifics of the relationship obtained as a result of the use of certain methods. [74]

4.1.4 Stage of the learning process

At this stage, the user is provided with content with a sequence based on the semantic proximity of the topics. Each user's content will be unique, because the initial knowledge and type of perception are taken into account. Thus, the learning path for users will be individual.

4.2 Information system architecture for education

Architecture introduces to large systems and how they are partitioned into subsystems and components, as well as how the structuring of these elements into a solution and the interfaces used to join them together facilitates communication and control. Different diagrams addresses challenges faced by software development organizations when developing large-scale software systems. The use of various notations is explored, with a focus on UML, and the role of architecture and detailed design specifications are considered from the perspective of risk management. [75] Use Case diagram is pictorial representation of a collection of tasks and users. Give an overview of how the system should interact with entities outside of the system. Use case diagram of intellectual information system in accordance with Figure 4.2.

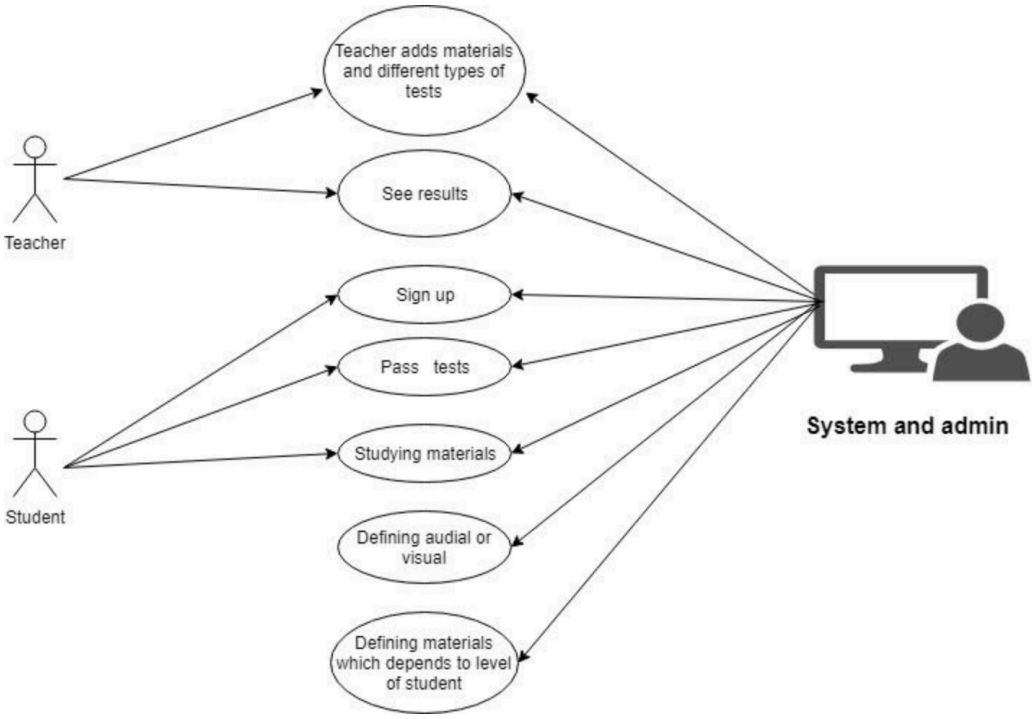

Figure 4.2 - Use-case diagram

Activity Diagrams describe how activities are coordinated to provide a service which can be at different levels of abstraction in accordance with Figure 4.3. [76]

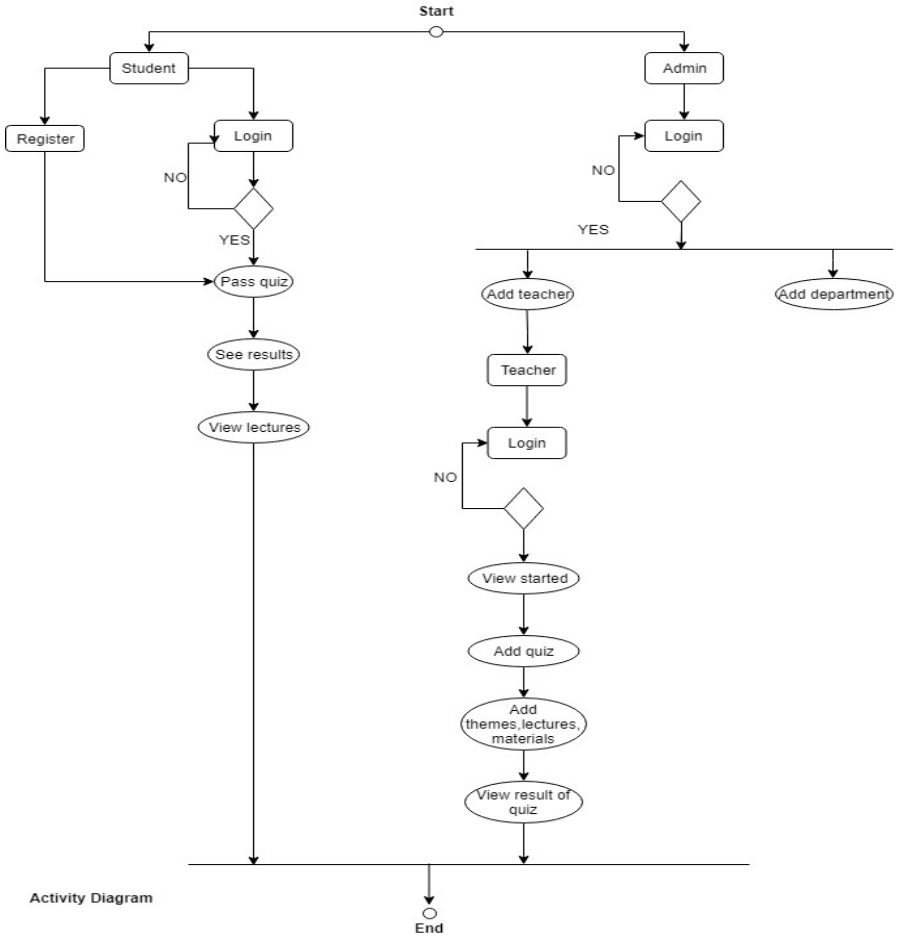

Figure 4.3 - Activity diagram

Sequence diagrams are used to represent how the system should respond to a request. Sequence diagrams show messages sent among objects between object “lifelines” – these are best for understanding interactions among objects. Sequence diagram of intellectual system in accordance with Figure 4.4 and Figure 4.5.

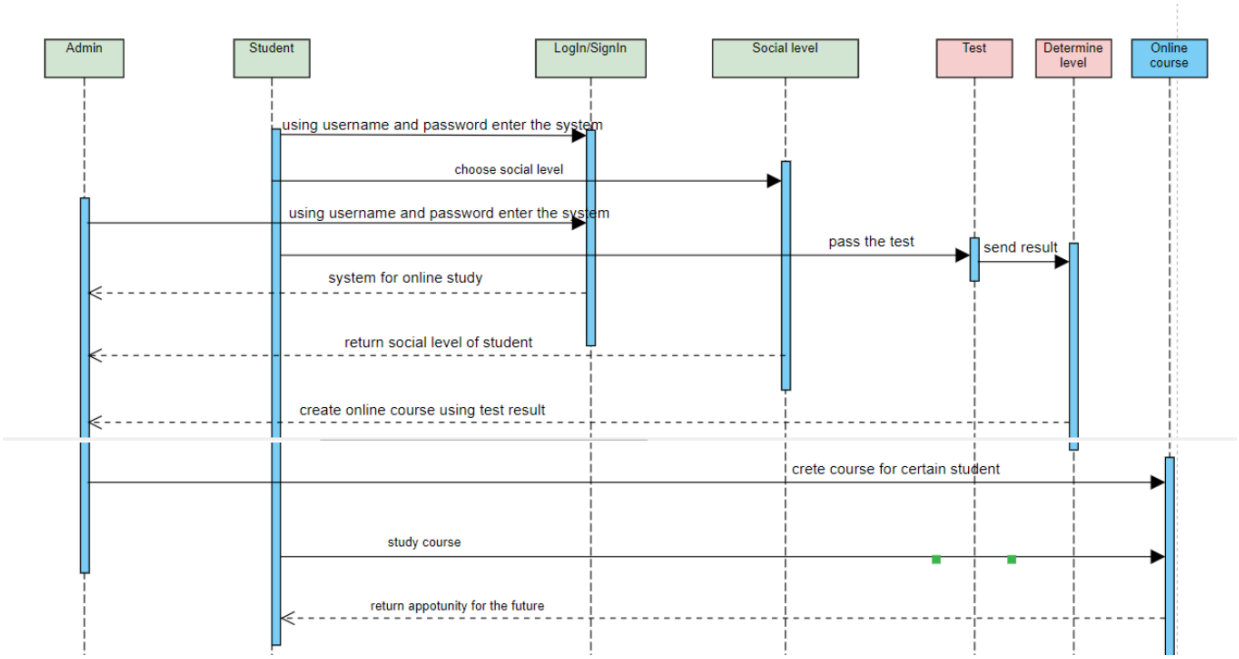

Figure 4.4 – Sequence diagram of system

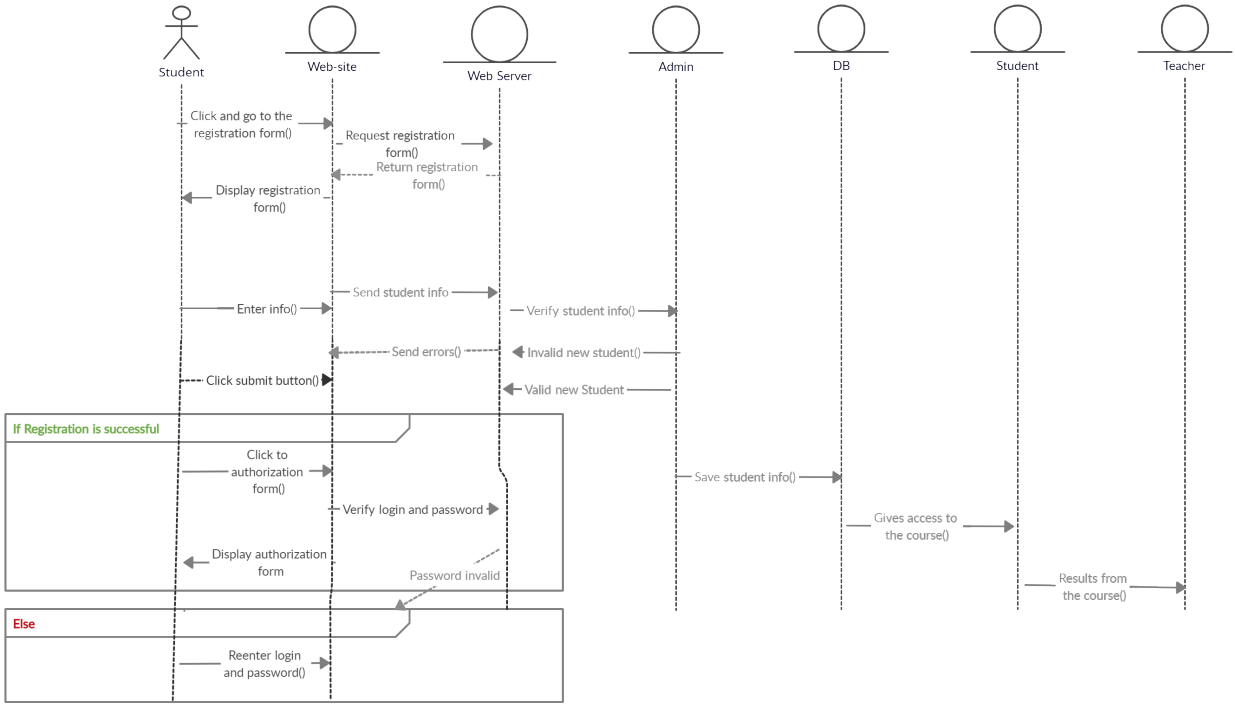

Figure 4.5 – Sequence diagram of system

An entity relationship diagram (ERD) shows the relationships of entity sets stored in a database. An entity in this context is an object, a component of data. An entity set is a collection of similar entities. These entities can have attributes that define its properties. ER diagram of intellectual system in accordance with Figure 4.6.

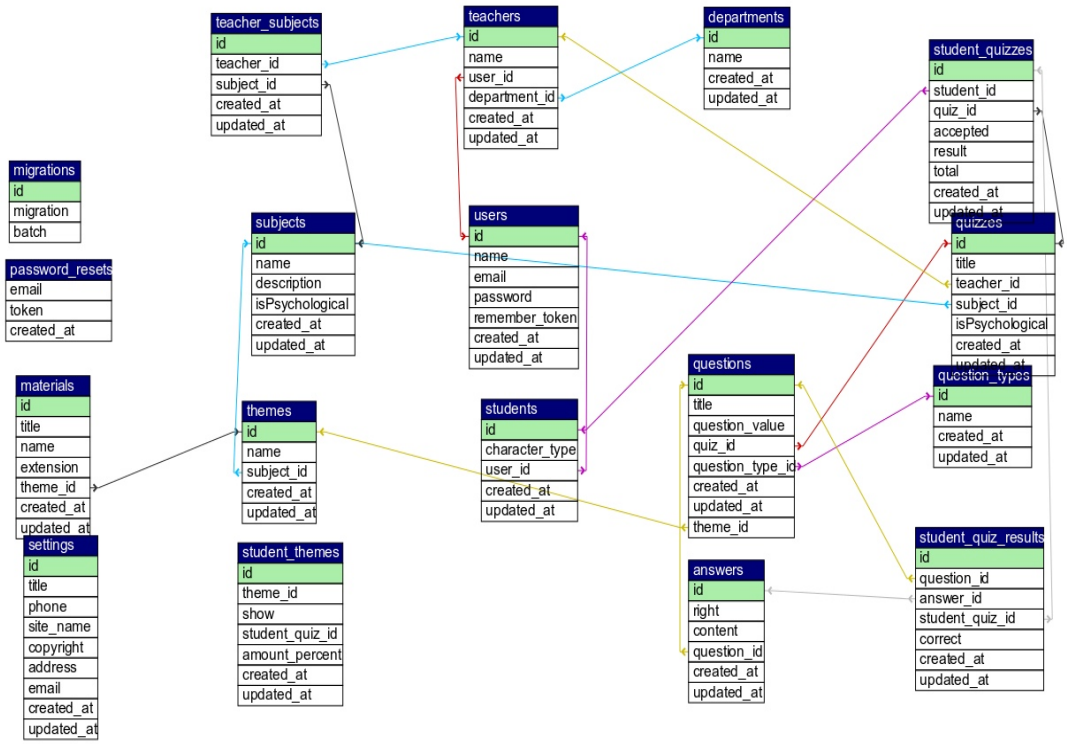

Figure 4.6 – Entity relationship diagram of system

Component Diagrams are used to show code modules of a system in Unified Modeling Language (UML). They are generally used for modeling subsystems. It represents how each and every component acts during execution and running of a system program. Component diagram of intellectual system in accordance with Figure 4.7-Figure 4.13. [77]

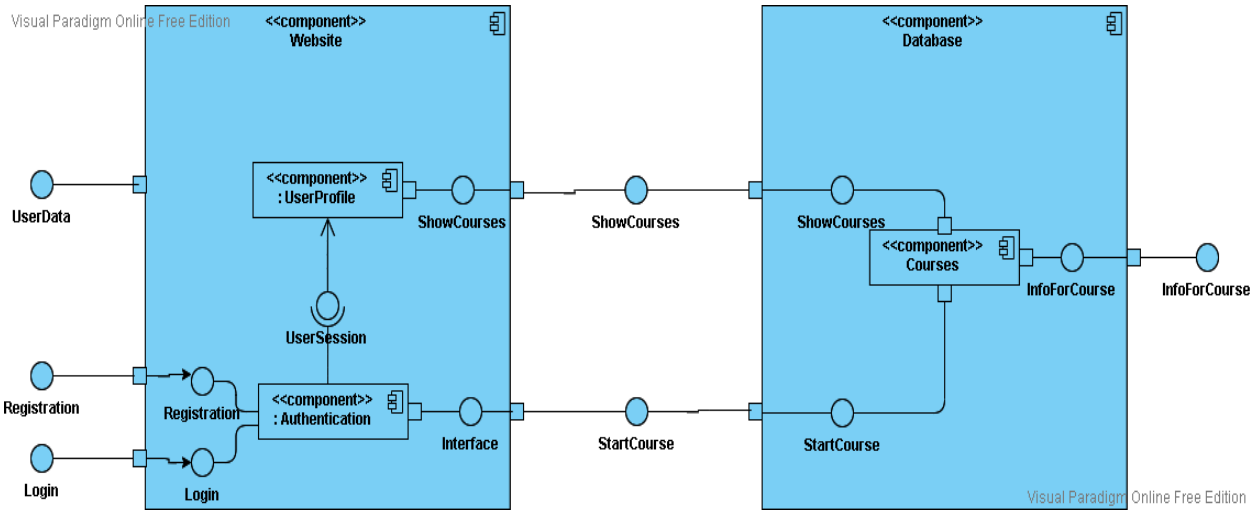

Figure 4.7 – component diagram of system

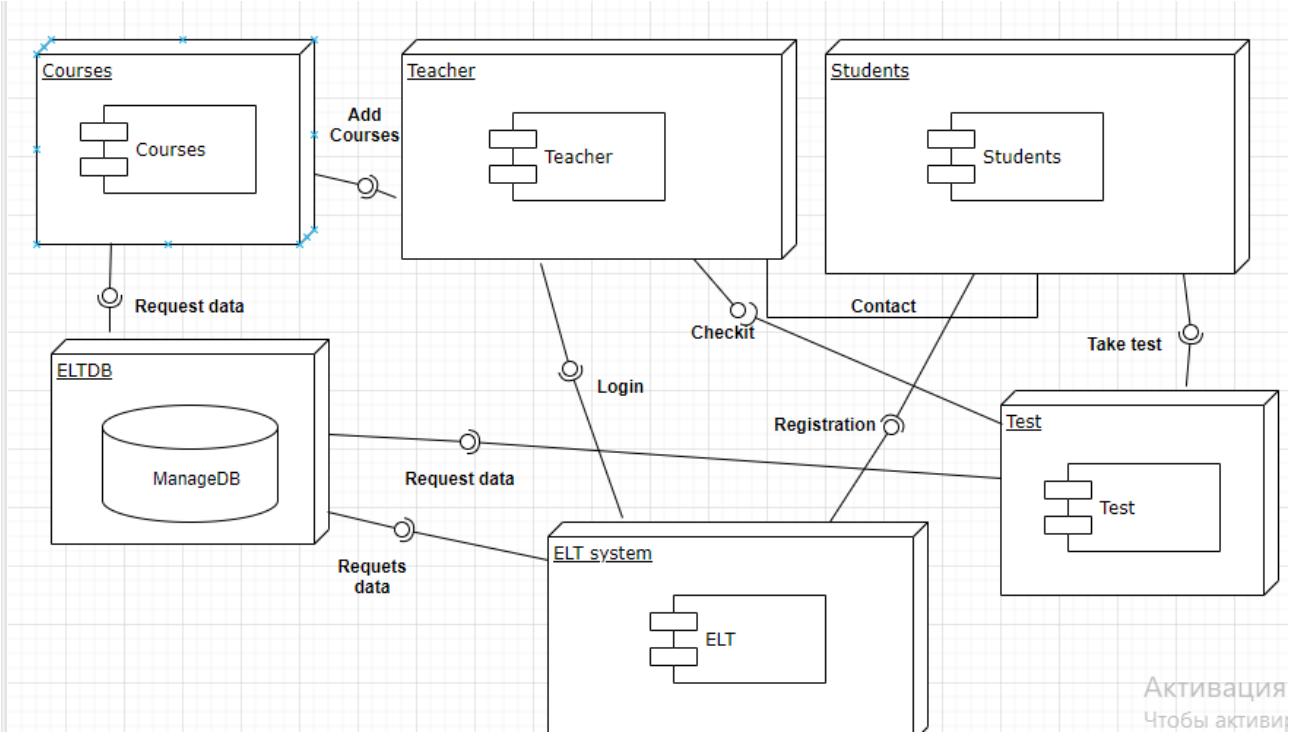

Figure 4.8 – component diagram of system

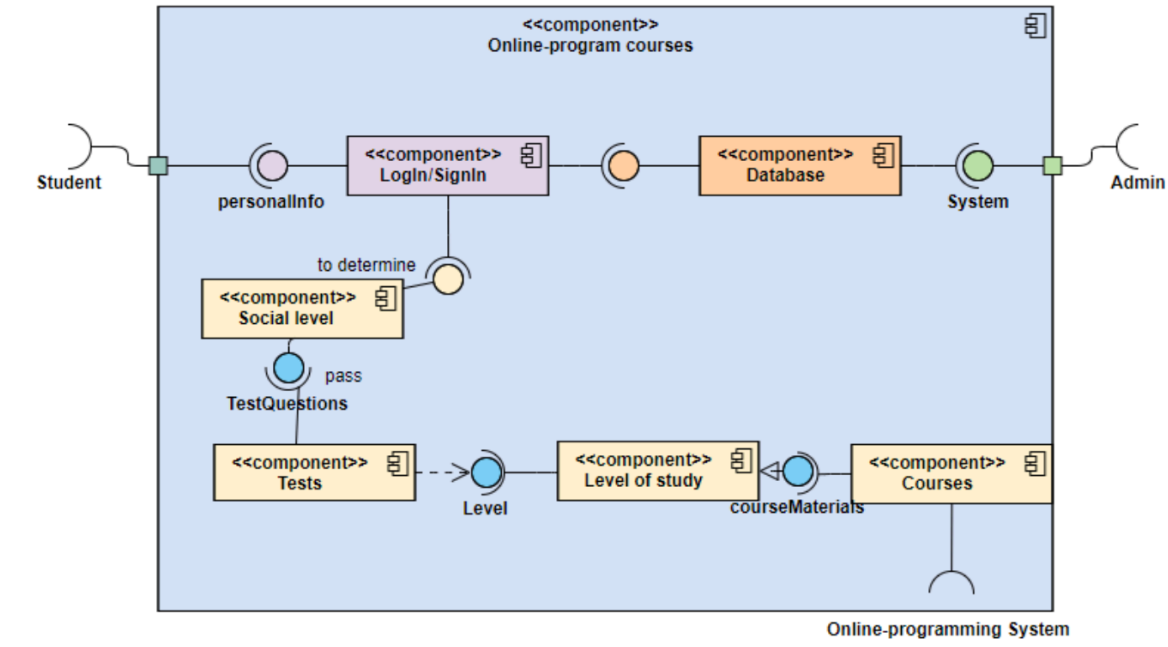

Figure 4.9 – component diagram of system

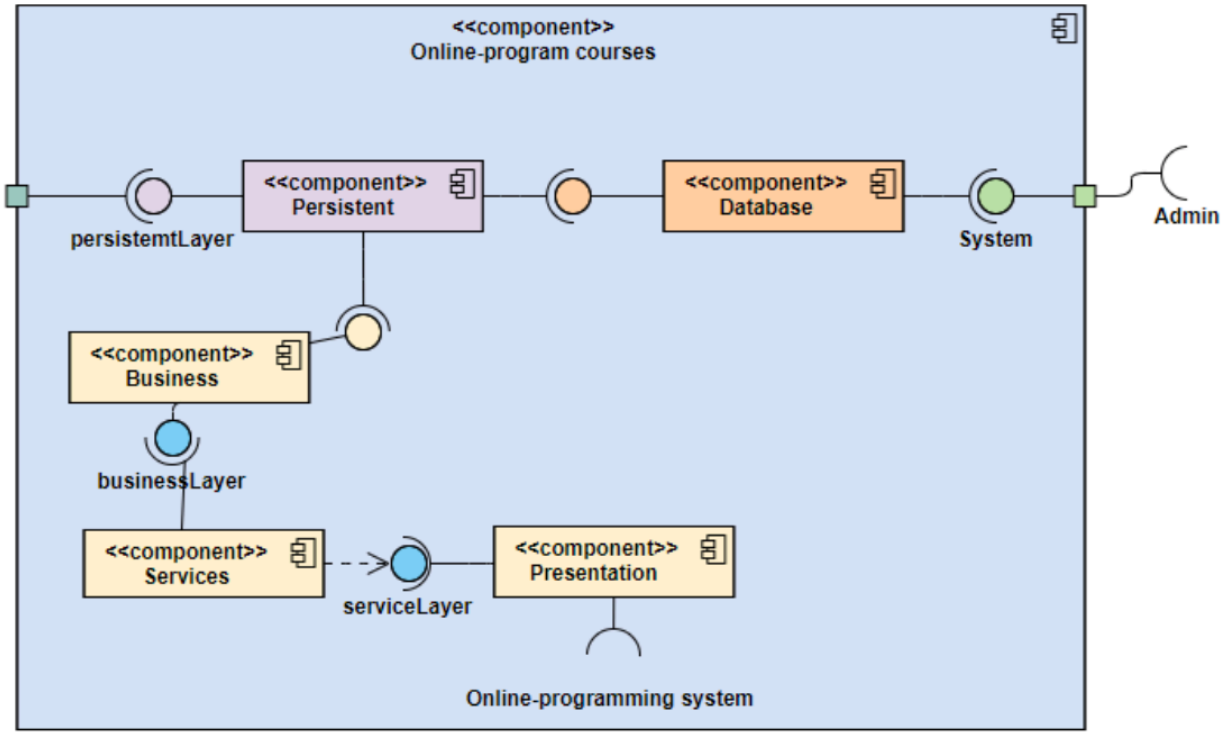

Figure 4.10 – component diagram of system

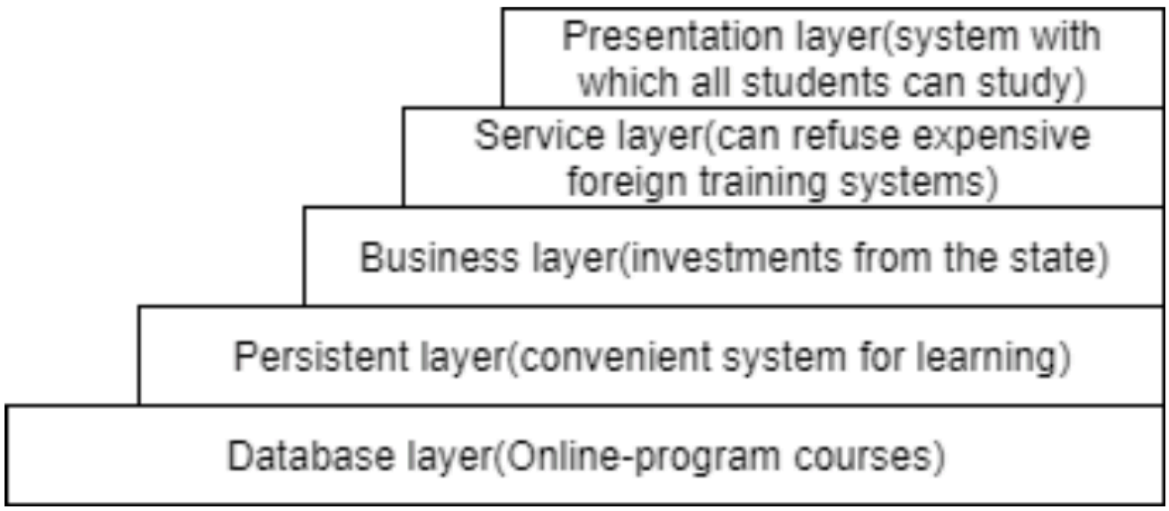

Figure 4.11 – component diagram with layered style

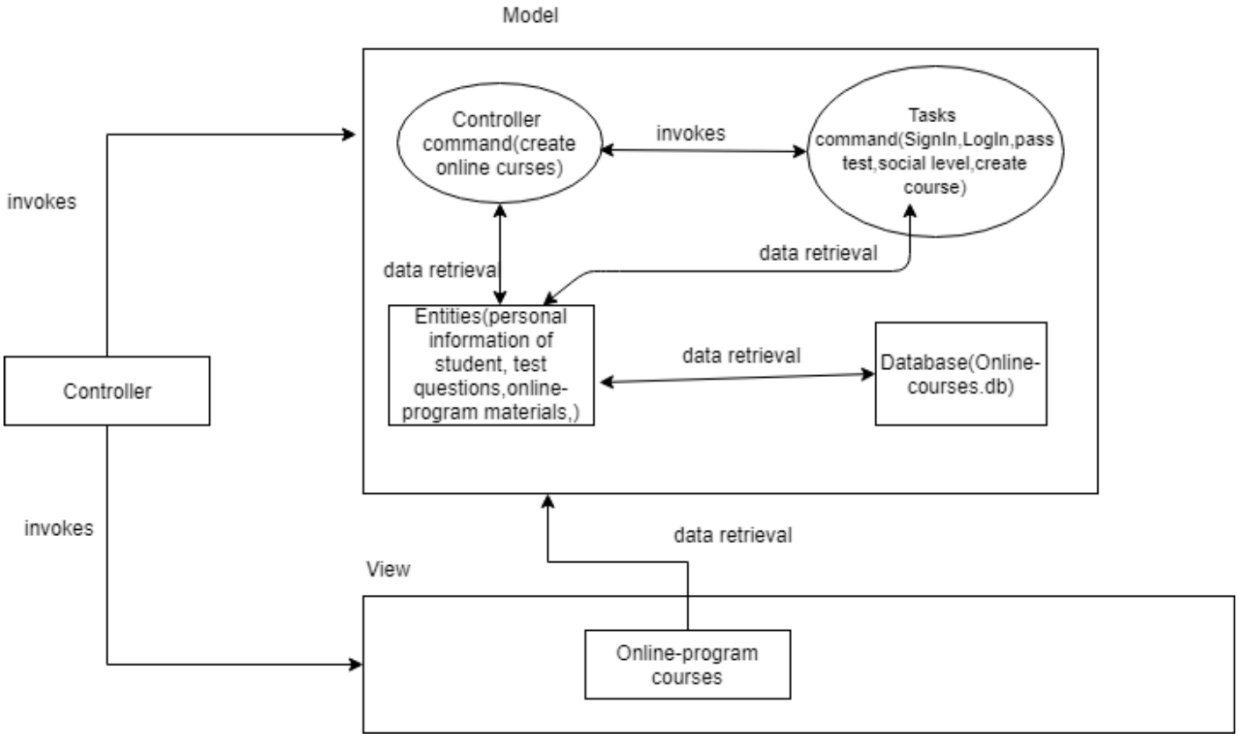

Figure 4.12 - Component diagram with model-centered style for system

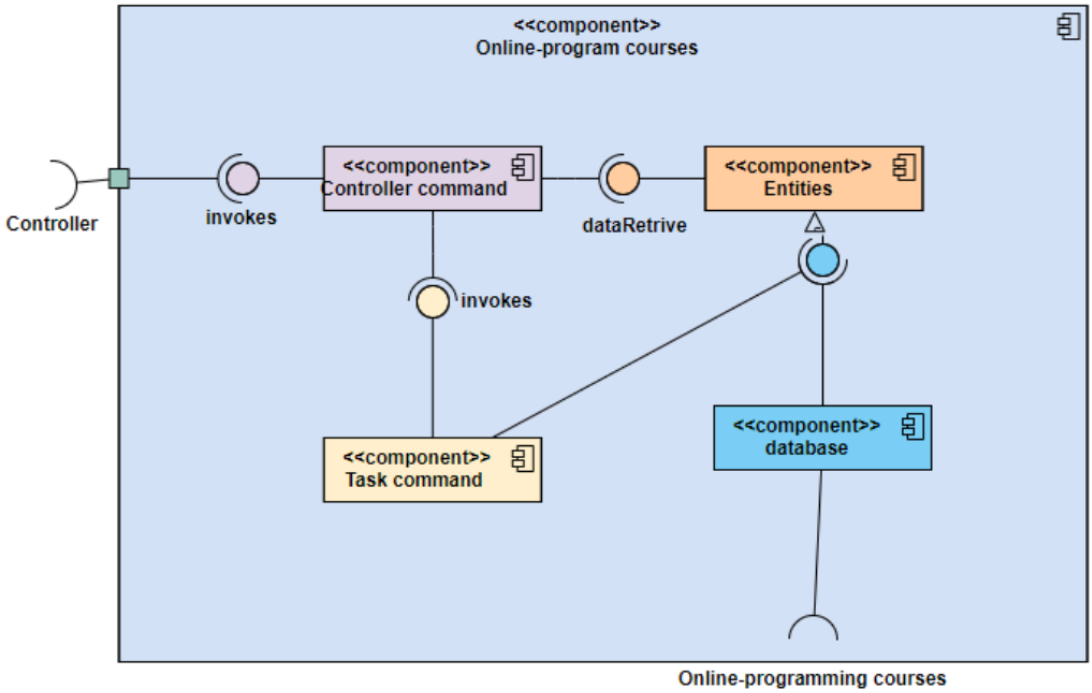

Figure 4.13 - Component diagram with model-centered style for system

4.3 Development of intellectual information and training system for education

The development of an intelligent information and training system is described edu-elt.kz. When visiting the system, users see the main page in accordance with Figure 4.14.

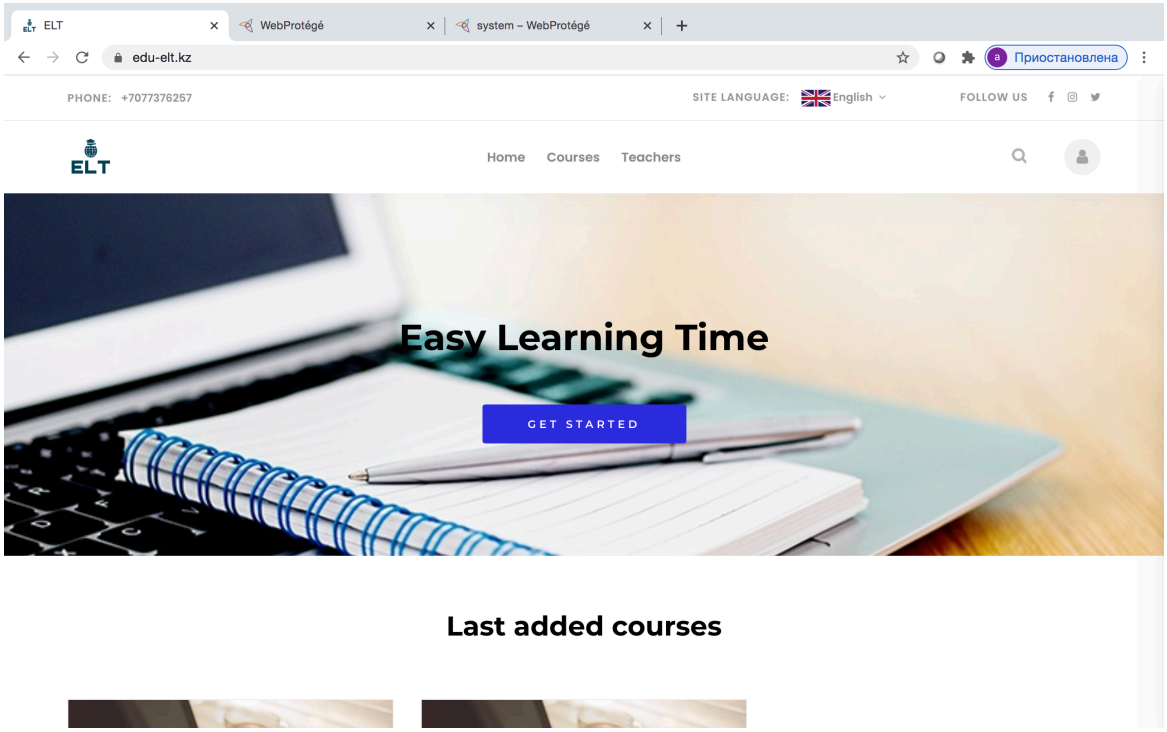

Figure 4.14 – Homepage of system

Below, the user can view all the courses in accordance with Figure 4.15.

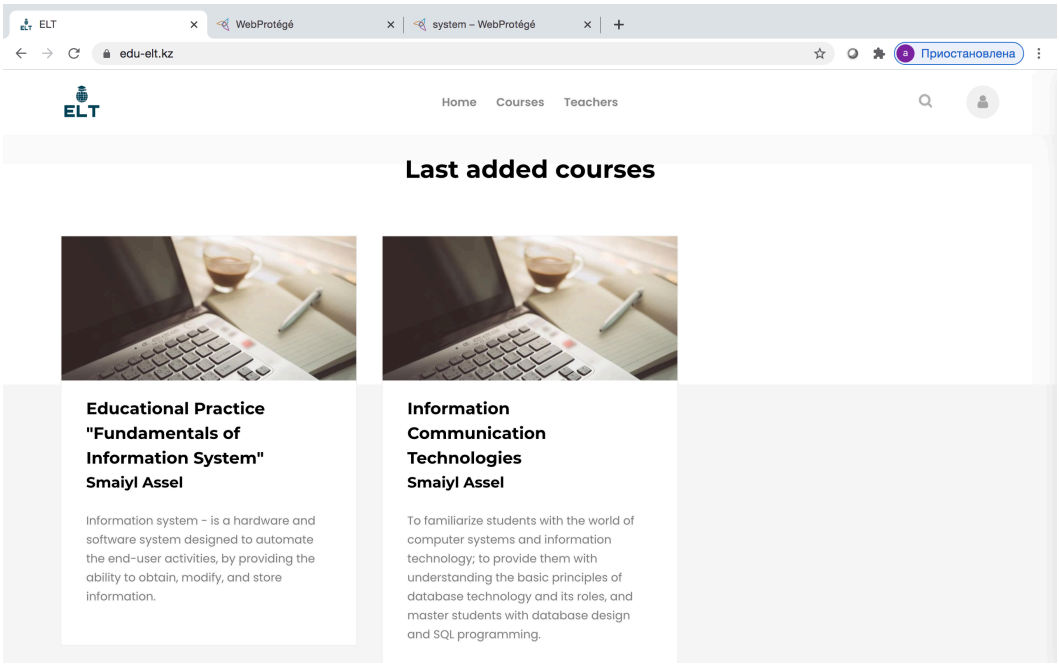

Figure 4.15 - Courses

The user can also view all the teachers in accordance with Figure 4.16.

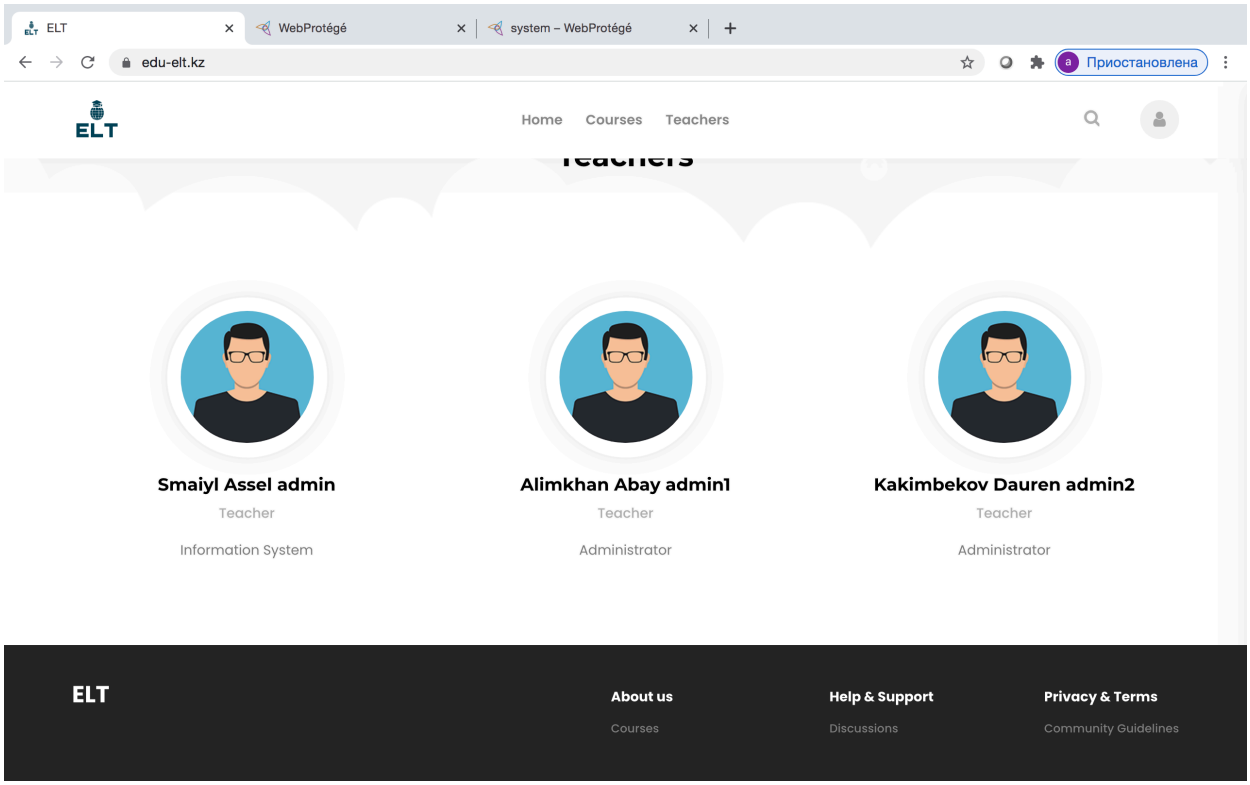

Figure 4.16 - Teachers

When logging in, the user can also register in the system, or log in, in accordance with Figure 4.17.

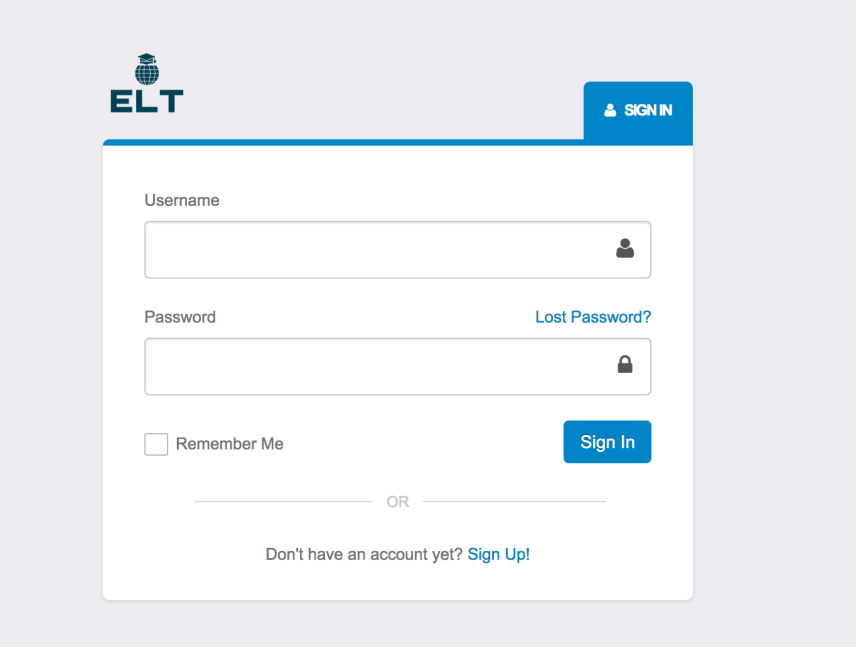

Figure 4.17 – Sign in

Registration / Sign up page in accordance with Figure 4.18:

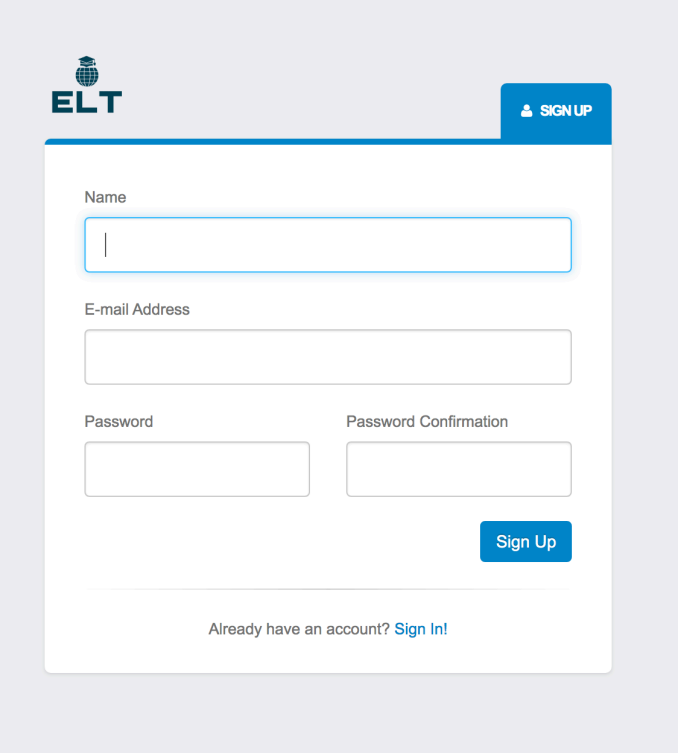

Figure 4.18 - Registration / Sign up page:

Next, the user sees Welcome page, in accordance with Figure 4.19:

Figure 4.19 – Welcome page

The user can view their profile and change it if necessary. At the entrance, the user can pass 2 tests: to determine the type of perception and to determine the level of initial knowledge. User can also configure the visibility of the test result in the profile in accordance with Figure 4.20 and Figure 4.21.

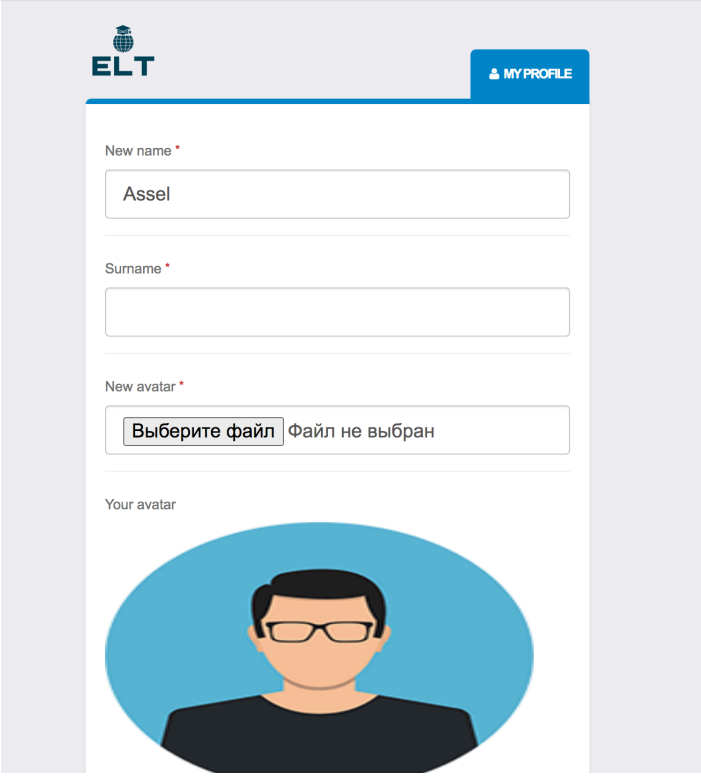

Figure 4.20 - My Profile page

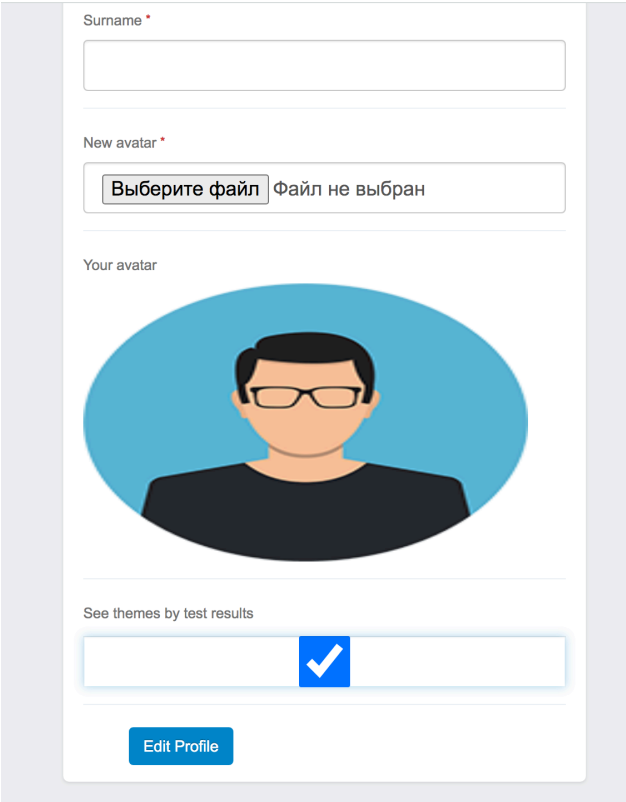

Figure 4.21 - Click see themes by test results

Questions for determining the type of perception are easy in accordance with Figure 4.22 below.

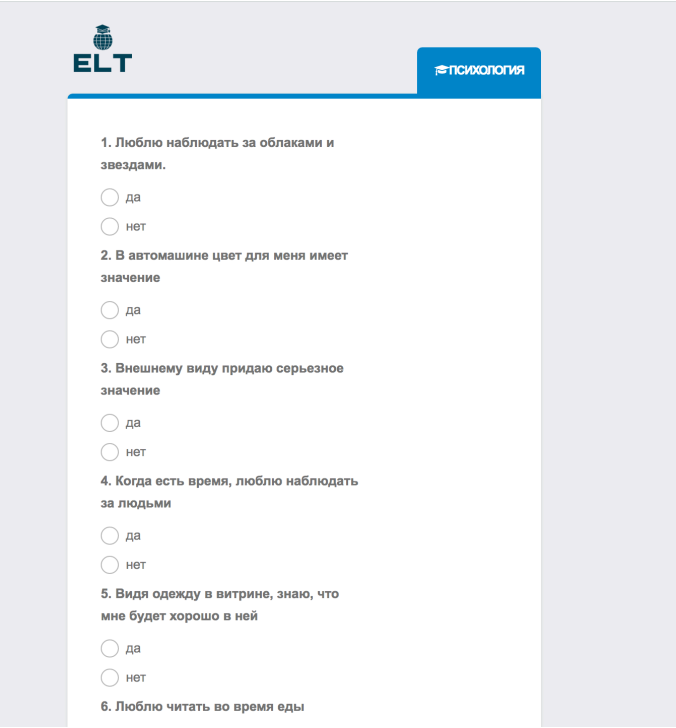

Figure 4.22 – Questions for defining type of perception

After that, if you click course «Educational Practice», opened new test for determine level of knowledge in accordance with Figure 4.23.

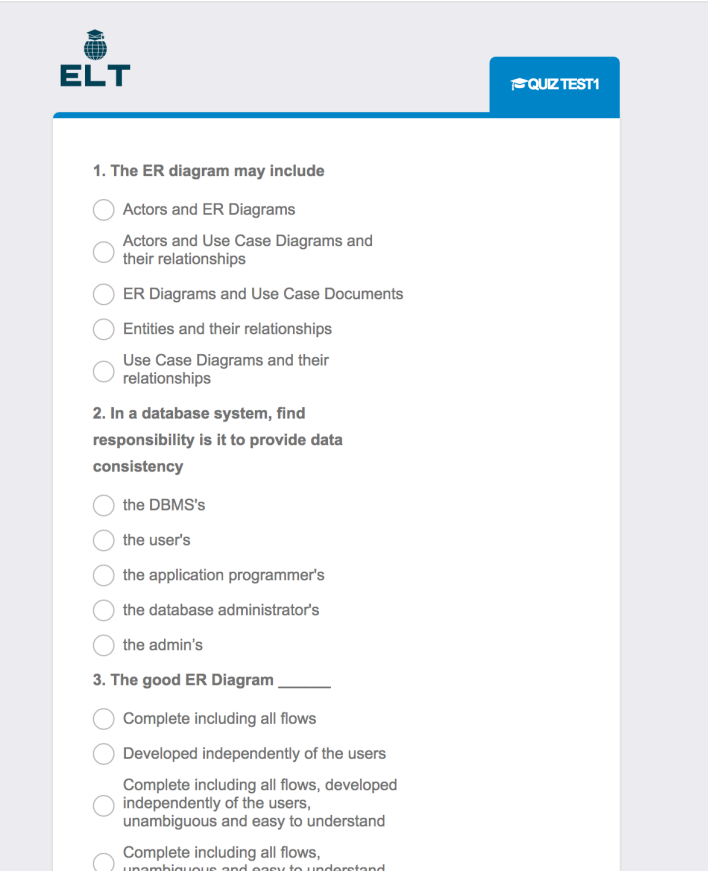

Figure 4.23 – Questions for defining initial knowledge

After completing the two tests, the user can start taking a specific course. In our case, Educational Practice “Fundamentals of Information System". After determining the initial knowledge, the system offers each user an individual learning path in accordance with Figure 4.24.

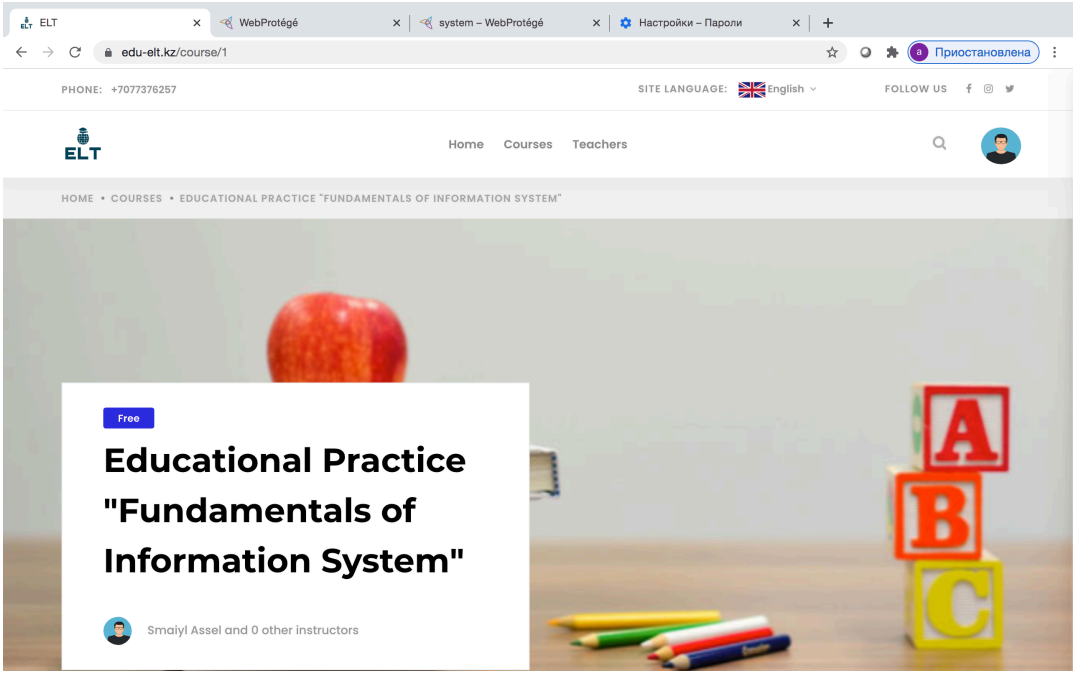

Figure 4.24 – Course Educational Practice “Fundamentals of Information System”

User can fee course features: duration of course, quantity of lectures, quizzes, test result and type of perception or character. In that page also user can see description of the course, indivial curriculum and members of the course in accordance with Figure 4.25 and Figure 4.26.

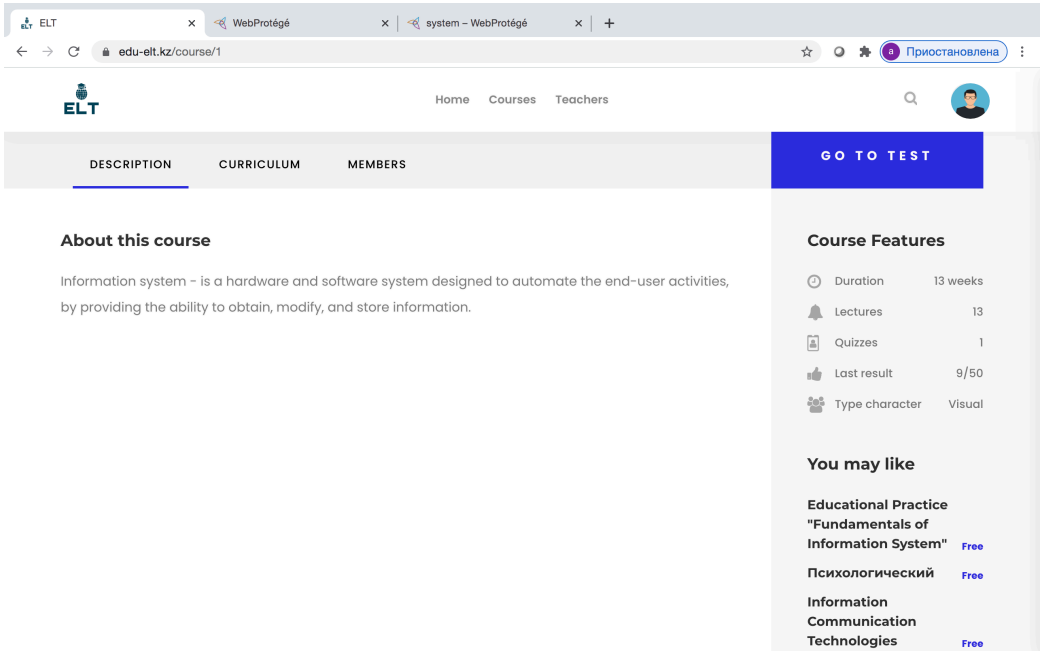

Figure 4.25 – Description of the course

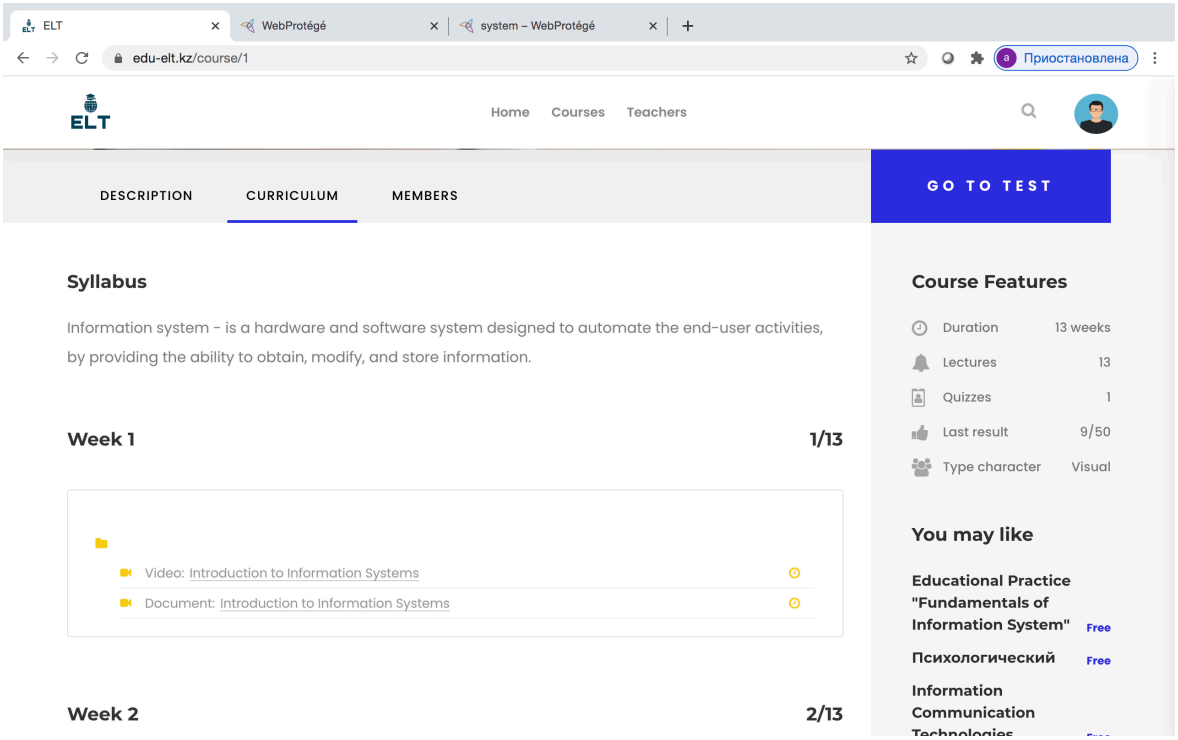

Figure 4.26 – Curriculum of the course

If user's type of character is visual, content should be video or animation graphic and document or presentation in accordance with Figure 4.27.

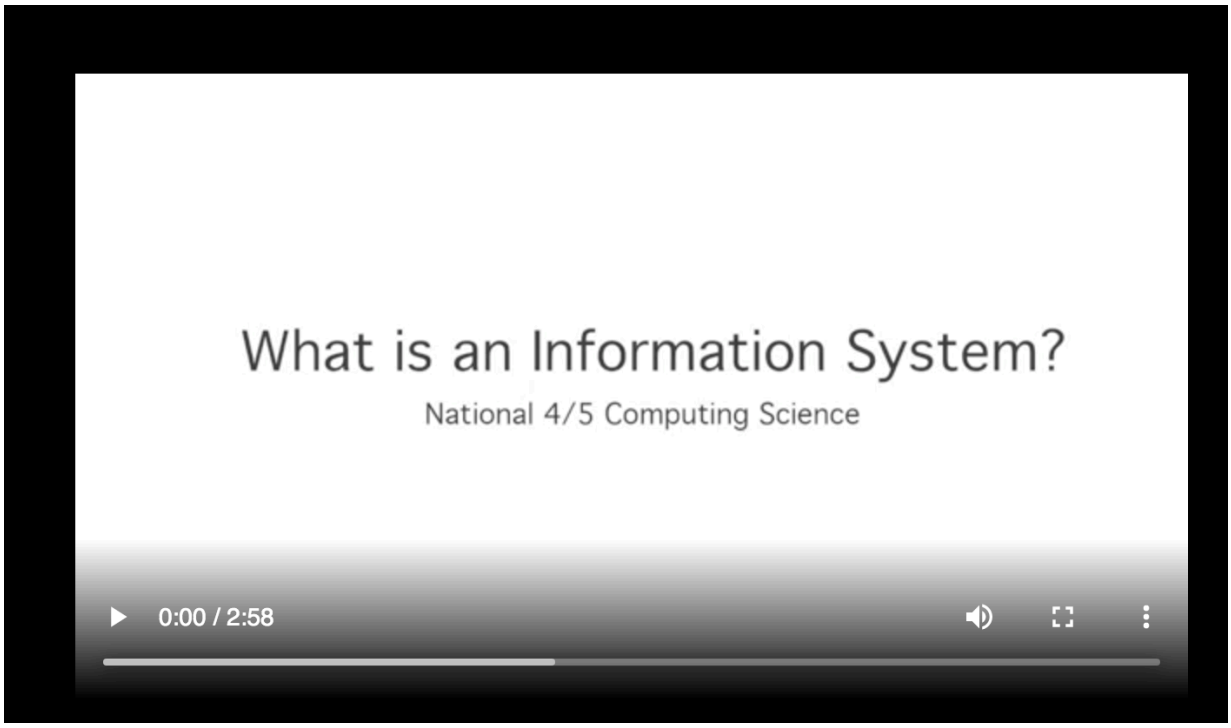

Figure 4.27 – content for visuals

If user's type of character is audial, content should be audio files and documents, which user can upload and open in computer in accordance with Figure 4.28.

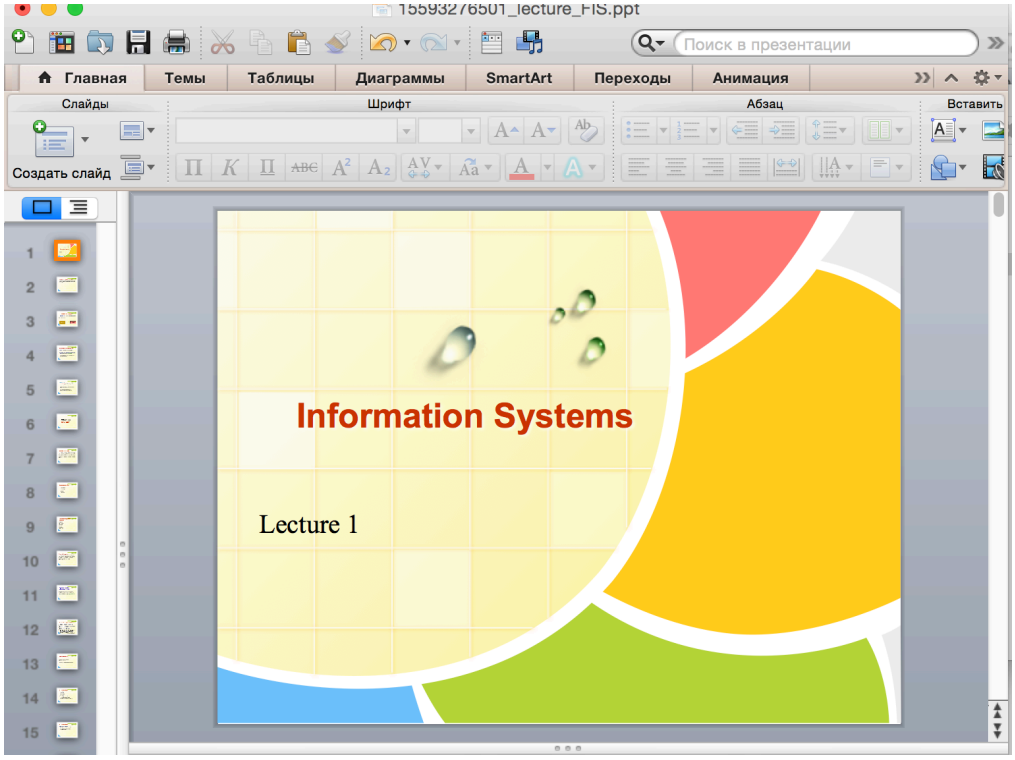

Figure 4.28 – presentation for lecture

System has admin panel for settings, manage all information, especially, upload content to the system, add new teachers, give privilegies, and etc. in accordance with Figure 4.29-4.31.

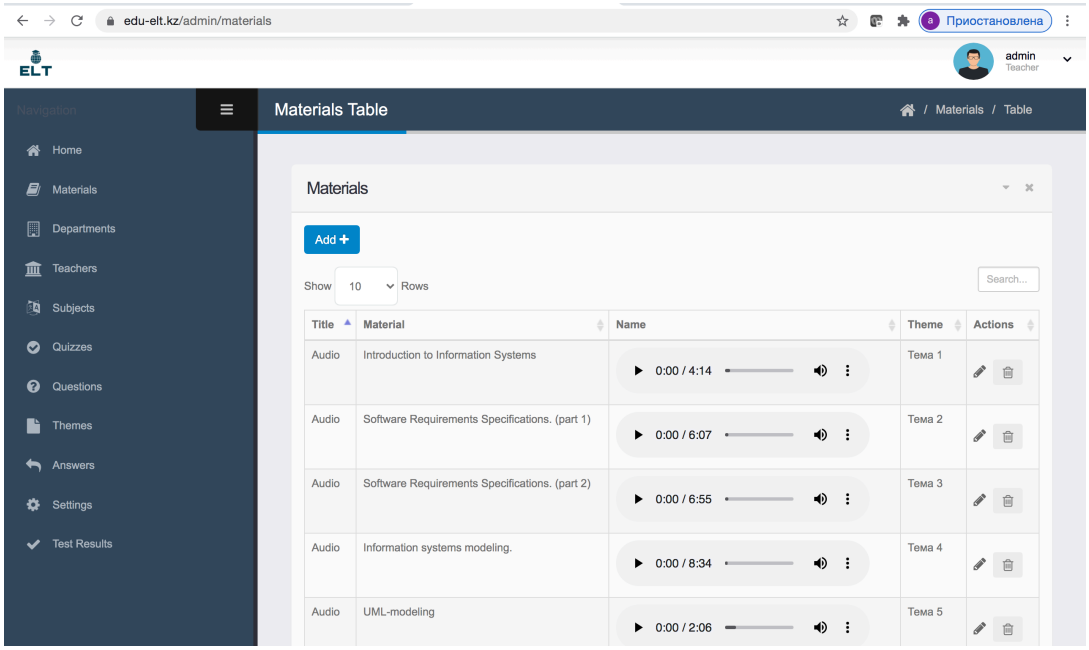

Figure 4.29 – Materials table

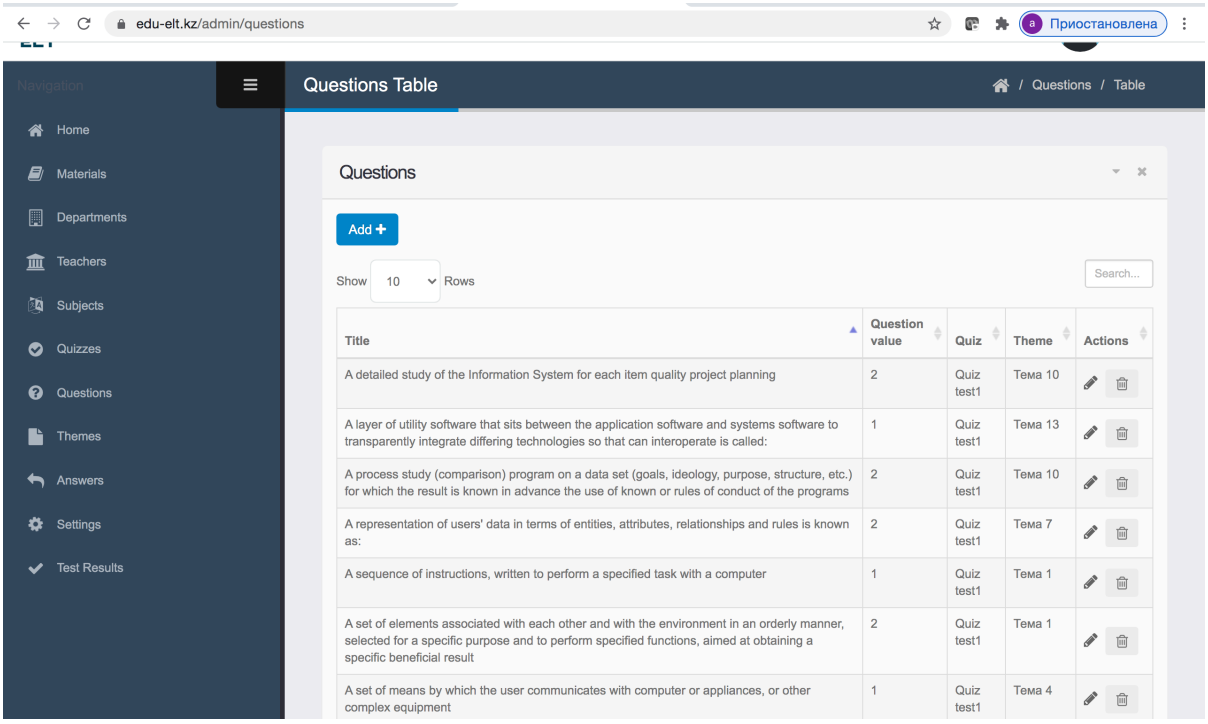

Figure 4.30 – Questions table

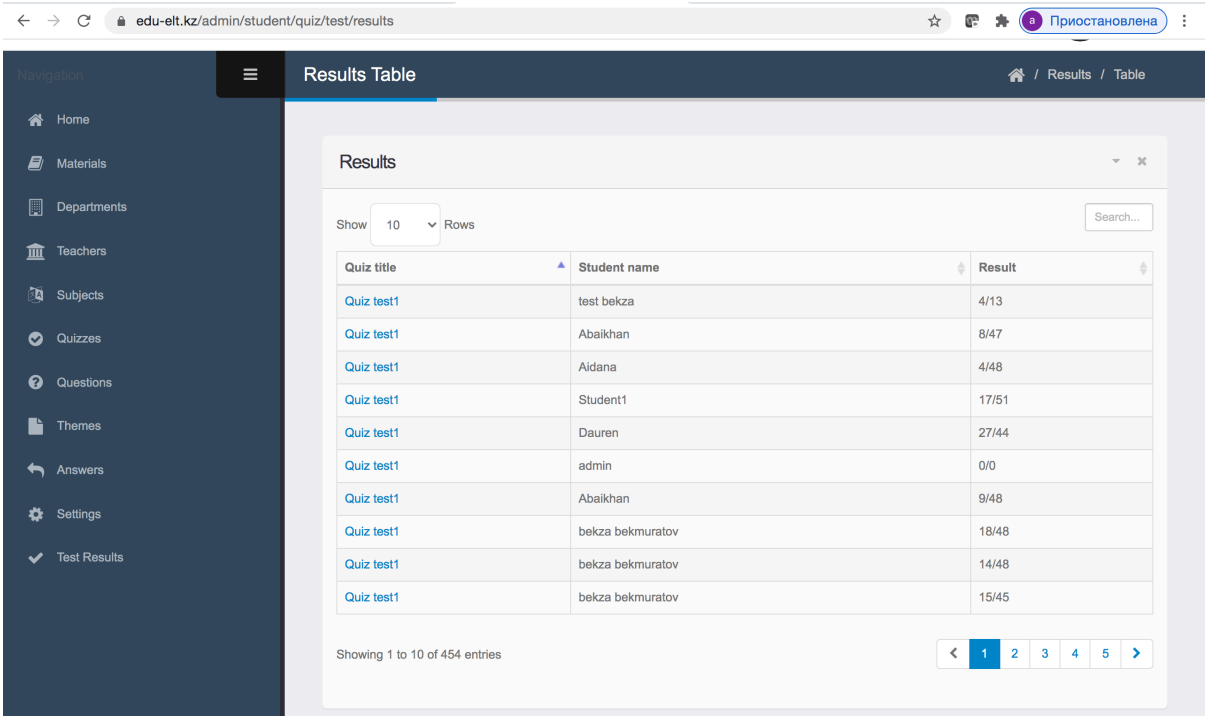

Figure 4.31 – Result table

Conclusions on the fourth section

SWOT analysis

S-strength	**W-weakness**
1. System adaptive learning by student characteristic 2. Adaptive search of web-service 3. Distance learning 4. 24/7 monitoring of students profit statistics 5. Self-updates 6. Reliable	1.Web-service has just one educational sphere
O-opportunity	**T-thread**
1. Implementation of web-service into other educational spheres 2. Online learning by video call	1. Refusal of the user from the adaptive learning quality improvement program

CONCLUSION

In this dissertation research, the development of an intelligent information and training system using models and methods of semantic analysis is carried out.

The theoretical research carried out within the framework of the dissertation allowed us to obtain the following results of practical and scientific significance:

1.The analysis and review of the existing systems of distance learning and adaptive intelligent systems. The requirements for the developed system are formulated.

2. Used scientific methods for the formation of content intended for any student, which takes into account the characteristics of the user.

3. Developed a methodology for building an educational resource based on semantic analysis.

4. A hypothesis is put forward about the semantic links between content-dependent and proved in the course of the study.

5. The sequence and management of the number of educational elements of an intelligent system for distance learning in the field of information technology are proposed. All components of this system are described in detail. A mathematical model is developed using semantic data analysis.

6. The result of the experimental study of the developed system, which demonstrates an increase in the quality of education and efficiency due to the semantic representation of resources, also creates conditions for the digitalization of the educational process.

The reliability of the theoretical results of the study is confirmed by scientific evidence, experimental studies, as well as experimental and industrial tests in real conditions.

Conducted of calculation of the degree of assimilation of educational material. It was found that in the experimental group the degree of assimilation increased by 0.62 points, and in the control group it fell by 0.13 points. The effectiveness of training with the use of IS is 1.15 times higher. The significance of differences between the control and experimental groups was assessed by a nonparametric method of statistical analysis using the Mann-Whitney test. The economic efficiency of the implemented work results is confirmed by the implementation act.

In the future works, can consider recommendations for using and creating new content for training and consider extracting knowledge from other types of data, such as audio and video, used in courses.

REFERENCES

1. Vostroknutov I. E. Homogeneity and aggressiveness of the visual environment in educational software. *Pedagogical Informatics*, 1997, no 4, pp. 43-50. (in Russian)
2. Zakharova T. B., Leshakov V. G. The role and place of information technologies in the modern course of informatics. *Tez.dokl. na Kongresse konferencii «Informacionnye tekhnologii v obrazovanii»,* [Abstract of the report at the congress of the conference "Information technologies in education], Moscow, 2002, pp. 27-31. (in Russian)
3. Kuandykov A., Serbin V., Smaiyl A. Role of big data technology in educational information systems. *Materialy Mezhdunarodnoj nauchno-prakticheskoj konferencii «Nauchno-innovacionnoe i social'no-ekonomicheskoe razvitie Kazahstana: novye koncepcii i sovremennye resheniya»*, [Materials of the International scientific and practical Conference "Scientific and innovative and socio-economic development of Kazakhstan: new concepts and modern solutions"], Shymkent, 2017, pp.235-236
4. Serbin V.V. *Tekhnologiya, metodologiya sozdaniya i razrabotka informacionno-obuchayushchih sistem* [Technology, methodology of creation and development of information and training systems]. Almaty, AiPET, 2010, 198 p. (In Russian)
5. Serbin V.V., Smaiyl A.M. Online courses on IT technologies for an information and training system. *Materialy Mezhdunarodnoj nauchno-prakticheskoj konferencii "Innovacii v obrazovanii i nauke", posvyashchennoj 25-letiyu Nezavisimosti Respubliki Kazahstan i 20-letiyu Universiteta Sulejmana Demirelya* [Materials of the International scientific-practical conference "Innovations in education and science" dedicated to the 25th anniversary of Independence of the Republic of Kazakhstan and 20th anniversary of Suleyman Demirel University], Kaskelen, 2016, pp. 87-91. (In Russian)
6. Smaiyl A.M. (2018). Overview of foreign educational information systems. Science and society – Methods and problems of practical application. Proceedings of the 5th International conference. Accent Graphics Communications & Publishing. Vancouver, Canada. 2018. pp. 7-11.
7. Serbin V.V., Syrymbayeva A., Tolebayeva K. Multi-criteria decision-making model for information learning system: a critique of the level of doubt. International Journal of eGovernance and Networks, 2015, no. 3, pp. 56–65.
8. Serbin V.V., Smirnova Yu.G. Psychometric indicator of doubt in computer test. Proc. Int. Research and Practice Conf. and I stage of the Championship in physical, mathematical and chemical sciences. London, 2013, pp. 43–46.
9. Serbin V. Methodology for measuring the level of users doubts: start of a new theory. European Applied Sciences is an international (no 1). Stuttgart, Germany: ORT Publishing. 2013. pp.230-233.

10. Serbin V. Multicriteria metod diagnosis of the intellectual and socialpsychological characteristics of personality. 2nd International Scientific Conference “European Applied Sciences: modern approaches in scientific researches”: Volume 3, Papers of the 1st International Scientific Conference (Volume 3). Stuttgart, Germany: Hosted by the ORT Publishing and The Center For Social and Political Studies “Premier”, 2013, pp.93-97.
11. Nogin V.D. Decisions in Multicriteria Environment: A Quantitative Approach. Moscow, Fizmatlit Publ., 2002, 176 p.
12. Serbin V. Technology, Methodology for the Creation and Development of Information and Learning Systems. Almaty, AIPET, 2010, 198 p.
13. Vinogradov G.P., Kuznetsov V.N. Modeling agent's behavior based subjective perceptions on the situation of choice. Scientific and Technical Information Processing, 2011, no. 3, pp. 58–72.
14. Teslinov A.G. Conceptual Design of Complex Solutions. St. Petersburg, Piter Publ., 2009, 288 p. (In Russian)
15. Hoenig C. The Problem Solving Journey: Your Guide To Making Decisions and Getting Results. Perseus Publishing, 2000, 283 p.
16. Ranyard R., Crozier W.R., Svenson O. Decision Making: Cognitive Models and Explanations. London, NY, Routledge, 1997.
17. Salas E., Klein G.A. Linking Expertise and Naturalistic Decision Making. Lawrance Erlbaum Associates, 2001, 464 p.
18. Stair R.M., Reynolds G. Principles of Information Systems: A Managerial Approach . Delmar Cengage Learning, 2007, 600 p.
19. Applegate L.M., Austin R.D., McFarlan F.W. Corporate Information Strategy and Management: Text and Cases. 7th ed. McGraw Hill, 2007, 657 p.
20. Robinson D. Aspect-Oriented Programming with the e Verification Language. Morgan Kaufmann, 2007, 265 p.
21. Serbin V.V., Syrymbayeva A.M. Research of multicriterial decision-making model for educational information systems. Scientific and Technical Journal of Information Technologies, Mechanics and Optics, 2016, vol. 16, no. 5, pp. 946–951. doi: 10.17586/2226- 1494-2016-16-5-946-951
22. M. West and S. Vosloo, *UNESCO Policy Guidelines for Mobile Learning*, 2015, https://unesdoc.unesco.org/ark:/48223/ pf0000219641.
23. Serbin V.V., Duisebekova K.S., Smaiyl A.M., Alimzhanova L.M., Zakirova G.D., Moldagulova A.N., Satybaldiyeva R.Zh., Maulenov Y.S. Multicreiteria model of students’ knowledge diagnostics based on the doubt measuring level method for E&M Learning. Mobile Information System, Volume 2019, Article ID 4260196, 11 pages, https://doi.org/10.1155/2019/4260196
24. Smaiyl A.M. Sequence of content generation for educational system. Bulletin of the Kazakh-British Technical University, vol. 16, no 3, 2019, pp. 309-317
25. Edagawa Y, Tani M, Sato T. Influence of active learning on academic achievement and difficulties in its implementation, WIAS Res Bull, 2015
26. Capuano N, Miranda S, Orciuoli F. Intelligent Collaborative e-Learning Systems and Applications, 2009

27. Kennedy J. Characteristics of Massive open online courses (MOOCs): A research review, 2009–2012, J Interact Online Learn, 2014
28. Ghadiri K, Qayoum MH, Junn E, Hsu P, Sujitparapitaya S. Revolution of learning caused by MOOC and flipped classes—challenges of San Jose State University, USA, 2013 JUCE J, 2013
29. Mori A. The path of university reform—The 1st flipped class, 2014
30. T. Tojo, O. Ono, N. B. M. Noh, and R. Yusof, "Interactive tutor robot for collaborative e-learning system," *Electrical Engi- neering in Japan*, vol. 203, no. 3, pp. 22–29, 2018
31. S. Joksimovic, O. Poquet, V. Kovanovic et al., "How do we model learning at scale? A systematic review of research on MOOCs," *Review of Educational Research*, vol. 88, no. 1, pp. 43–86, 2018
32. M. Firat, H. Kilinc, and T. V. Yuzer, "Level of intrinsic motivation of distance education students in e-learning en- vironments," *Journal of Computer Assisted Learning*, vol. 34, no. 1, pp. 63–70, 2018.
33. M. Farhan, S. Jabbar, M. Aslam et al., "IoT-based students interaction framework using attention-scoring assessment in e-learning," *Future Generation Computer Systems*, vol. 79, no. 3, pp. 909–919, 2018.
34. S. B. Eom and N. J. Ashill, "A system's view of e-learning success model," *Decision Sciences Journal of Innovative Ed- ucation*, vol. 16, no. 1, pp. 42–76, 2018.
35. S. Bhattacharya, S. Roy, and S. Chowdhury, "A neural network-based intelligent cognitive state recognizer for confidence-based e-learning system," *Neural Computing and Applications*, vol. 29, no. 1, pp. 205–219, 2018.
36. M. J. Marcelino, T. Pessoa, C. Vieira, T. Salvador, and A. J. Mendes, "Learning computational thinking and scratch at distance," *Computers in Human Behavior*, vol. 80, pp. 470–477, 2018.
37. P. Debiec, "Effective learner-centered approach for teaching an introductory digital systems course," *IEEE Transactions on Education*, vol. 61, no. 1, pp. 38–45, 2018.
38. S. Gul, I. Mahajan, H. Shafiq, M. Shafi, and T. A. Shah, "Massive open online courses: hype and hope," *DESIDOC Journal of Library & Information Technology*, vol. 38, no. 1, pp. 63–66, 2018.
39. S. Farid, R. Ahmad, M. Alam, A. Akbar, and V. Chang, "A sustainable quality assessment model for the information delivery in e-learning systems," *Information Discovery and Delivery*, vol. 46, no. 1, pp. 1–25, 2018.
40. V. Bradac and B. Walek, "A comprehensive adaptive system for e-learning of foreign languages," *Expert Systems with Applications*, vol. 90, pp. 414–426, 2017.
41. Y. Ren, Z.-X. Dai, X.-H. Zhao, M.-M. Fei, and W.-T. Gan, "Exploring an on- line course applicability assessment to assist learners in course selection and learning effectiveness im- proving in e-learning," *Learning and Individual Differences*, vol. 60, pp. 56–62, 2017.
42. A. Cohen, "Analysis of student activity in web-supported courses as a tool for

predicting dropout," *Educational Tech- nology Research and Development*, vol. 65, no. 5, pp. 1285– 1304, 2017.
43. M. Hamada and M. Hassan, "An enhanced learning style index: implementation and integration into an intelligent and adaptive e-learning system," *Eurasia Journal of Mathematics, Science and Technology Education*, vol. 13, no. 8, pp. 4449– 4470, 2017.
44. E. Levina, A. Masalimova, N. Kryukova et al., "Structure and content of e-learning information environment based on geo- information technologies," *Eurasia Journal of Mathematics, Science and Technology Education*, vol. 13, no. 8, pp. 5019– 5031, 2017.
45. S. Sweta and K. Lal, "Personalized adaptive learner model in E-learning system using FCM and fuzzy inference system," *International Journal of Fuzzy Systems*, vol. 19, no. 4, pp. 1249–1260, 2017.
46. A. Bralic and B. Divjak, "Integrating MOOCs in traditionally taught courses: achieving learning outcomes with blended learning," *International Journal of Educational Technology in Higher Education*, vol. 15, pp. 1–16, 2018.
47. P. Gregori, V. Mart ınez, and J. J. Moyano-Ferna ndez, "Basic actions to reduce dropout rates in distance learning," *Eval- uation and Program Planning*, vol. 66, pp. 48–52, 2018.
48. J. G. Quintana and J. M. Perez, "The students empowerment in the sMOOC," *Revista Complutense de Educacion*, vol. 29, no. 1, pp. 43–60, 2018.
49. Tang, Shujun, Learning Mechanism and Function Characteristics of MOOC in the Process of Higher Education, EURASIA JOURNAL OF MATHEMATICS SCIENCE AND TECHNOLOGY EDUCATION, 2017, vol. 13, no 12, pp. 8067-8072
50. V. Serbin and Y. Gorbunov, "Analysis of MPP-systems stress testing based on big data," *International Journal of Advances in Electronics and Computer Science*, vol. 4, no. 2, pp. 56–60, 2017.
51. V. V. Serbin and A. M. Syrymbayeva, "Research of multi- criterial decision-making model for educational information systems," *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, vol. 16, no. 5, pp. 946– 951, 2016.
52. Nauchnaya jelektronnaya biblioteka Kiberleninka [In Russian: Scientific electronic library CyberLeninka] [Elecronic resurce]. –URL: https://cyberleninka.ru/article/n/informatsionno-obuchayuschaya-sreda-problemy-formirovaniya-i-organizatsii-uchebnogo-protsessa (Date of application 03.10.2018).
53. Geoff Petty. Teaching Today A Practical Guide. OUP Oxford Publ, 4th Revised edition, 2009, 23 p.
54. Fetiskin N.P., Kozlov V.V., Manujlov G.M. Diagnostika dominirujushhej perceptivnoj modal'nosti [In Russian: Diagnosis of the dominant perceptual modality] /. – M., 2002. p.237-238
55. Garrett Grolemund, Hadley Wickham. R for Data Science. O'Reilly Media Publ, 1 edition, 2017, 432 p.

56. Smaiyl A.M. Data visualization based on the R Language for the information and training system. The Bulletin of Kazakh Academy of Transport and Communications named after M.Tynyshpayev 2018, Vol. 107, No. 4, pp. 216-223
57. Aliaskarov S., Saimassayeva Sh., Smaiyl A. «International Journal of Future Computer and Communication (IJFCC)» Singapore, 2019, vol. 8, no. 3, pp. 85-88. doi: 10.18178/ijfcc.2019.8.3.545
58. Kozlov, A. N. Intelligent information systems: textbook /A. N. Kozlov; Min-vo s-X. Russian Federation, FGBOU VPO Perm State Agricultural Academy. - Perm: Publishing house of the Perm State Agricultural Academy, 2013. - 278 p
59. A. V. Shishin and O. R. Kozyriov, "using ontology in the system of semantic analysis of cases," in *Bussines-Informatics*, vol. III, 05 - 2008, pp. 25–28.
60. V. V. Nikitayev, "ONTOLOGIES AND PROBLEMS OF TECHNICAL REALITY," in *Humanitarian Portal*, 2009.
61. G. A. Dodder, "Becision-making methods and systems," textbook Krasnoyarsk Siberian Federal University, 2016
62. Smaiyl A. Semantic representation of data in an intellectual information system. Bulletin of the Almaty University of Energy and Communications, vol. 3, no 50, 2020
63. Korobkin, A.A. (2009), Development of models and methods of decision making using artificial intelligence for the organization of the educational process [Razrabotka modelei I metodov prinyatiya resheniy s primeneniem iskusstvennogo intellekta dlya organizacii uchebnogo processa; dis. kand. tech. nauk], Voronezh, 26 p.
64. Christophiles,A.A. (1987), Graph theory.Algorithmic approach [Teoriya graphov.Algoritmicheskyi podhod], Mir, Moscow, 432p.
65. Krivossheyev, A.O. (2006), Development and use of computer tutorials, 2nd ed. [Razrabotka I ispolzovanie kompiuternyh program, 2oe izd.], Information technologies, 14 p.
66. D. V. Bondarchuk, "CALCULATING THE SEMANTIC RELATEDNESS OF TERMS WITH THE CONTEXT SET," *J. Name Stand. Abbrev.*, in *Informatics computer engineering and management.* 2017. pp 172-179
67. Armstrong RA. Should Pearson's correlation coefficient be avoided? Ophthalmic Physiol Opt 2019; 39: 316–327.
68. James Lewis, Stephan Ossowski, Justin Hicks, Mounir Errami, Harold R. Garner, Text similarity: an alternative way to search MEDLINE, Bioinformatics, Volume 22, Issue 18, 15 September 2006, Pages 2298–2304
69. Fletcher and Islam, Australasian Journal of Information Systems, 2018, Vol 22
70. Turdakov D., Velikhov P. 2008. Semantic relatedness metric for wikipedia concepts based on link analysis and its application to word sense

disambiguation. 5th Spring Young Researchers Colloquium on Databases and Information Systems, SYRCoDIS'2008, Proceedings.
71. A. Smaiyl, M. Brambilla, A. Kuandykov and Z. Sarsenova, "Methods for an Intellectual Information Learning System," *2021 IEEE International Conference on Smart Information Systems and Technologies (SIST)*, 2021, pp. 1-4, doi: 10.1109/SIST50301.2021.9465898
72. Available at: https://www.ispringsolutions.com/blog/what-is-microlearning, (last visited on June 24, 2021)
73. Starichenko B. E. Processing and presentation of pedagogical research data using a computer, Ural. gos.ped. un-T. Ekaterinburg, 2004. – 218 p.
74. Makarenko S. I. Intellectual information systems: a tutorial - Stavropol: SF MGGU after M. A. Sholokhova, 2009 .-- 206 p .: ill.
75. Technical report: Software Architecture: An Executive Overview: http://www.rai.com/, last visited on May 30, (2021)
76. Available at: https://www.visual-paradigm.com/guide/uml-unified-modeling-language/what-is-activity-diagram/, last visited on May 29, (2021)
77. Available at: https://www.geeksforgeeks.org/component-based-diagram/, last visited on May 25, (2021)

APPENDIX A - SOURCE CODE

User.php

```php
<?php

namespace App;

use Illuminate\Notifications\Notifiable;
use Illuminate\Foundation\Auth\User as Authenticatable;

class User extends Authenticatable
{
    use Notifiable;

    /**
     * The attributes that are mass assignable.
     *
     * @var array
     */
    protected $fillable = [
        'name', 'email', 'password',
    ];

    public function teacher(){
        return $this->hasOne('App\Teacher');
    }

    public function student(){
        return $this->hasOne('App\Student');
    }

    /**
     * The attributes that should be hidden for arrays.
     *
     * @var array
     */
    protected $hidden = [
        'password', 'remember_token',
```

```
    ];

    public function  profile(){
        return $this->hasOne('App\Profile');
    }
}
```

Student.php

```
<?php

namespace App;

use Illuminate\Database\Eloquent\Model;

class Student extends Model
{
    protected $fillable = ['character_type', 'user_id'];

    public function user(){
        return $this->belongsTo('App\User');
    }

    public function studentQuizzes(){
        return $this->hasMany('App\StudentQuiz');
    }
}
```

StudentQuiz.php

```
<?php

namespace App;
```

```php
use Illuminate\Database\Eloquent\Model;

class StudentQuiz extends Model
{
    protected $fillable = ['student_id', 'quiz_id',
'accepted', 'result', 'total'];

    public function quiz(){
        return $this->belongsTo('App\Quiz');
    }

    public function student(){
        return $this->belongsTo('App\Student');
    }

    public function studentQuizResults(){
        return $this->hasMany('App\StudentQuizResult');
    }
}
```

StudentQuizResult.php

```php
<?php

namespace App;

use Illuminate\Database\Eloquent\Model;

class StudentQuizResult extends Model
{
    protected $fillable = ['question_id', 'answer_id',
'correct', 'student_quiz_id'];

    public function question(){
```

```
        return $this->belongsTo('App\Question');
    }

    public function answer(){
        return $this->belongsTo('App\Answer');
    }

    public function studentQuiz(){
        return $this->belongsTo('App\StudentQuiz');
    }
}
```

Teacher.php

```
<?php

namespace App;

use Illuminate\Database\Eloquent\Model;

class Teacher extends Model
{

    protected $fillable = ['name', 'user_id',
'department_id',];

    public function user(){
        return $this->belongsTo('App\User');
    }

    public function department(){
        return $this->belongsTo('App\Department');
    }

    public function quizzes(){
        return $this->hasMany('App\Quiz');
    }
```

```
    public function subjects(){
        return $this->belongsToMany('App\Subject',
'teacher_subjects', 'teacher_id', 'subject_id');
    }
}
```

```
```

package.json

```
{
    "private": true,
    "scripts": {
        "dev": "npm run development",
        "development": "cross-env NODE_ENV=development
node_modules/webpack/bin/webpack.js --progress --hide-
modules --config=node_modules/laravel-
mix/setup/webpack.config.js",
        "watch": "cross-env NODE_ENV=development
node_modules/webpack/bin/webpack.js --watch --progress --
hide-modules --config=node_modules/laravel-
mix/setup/webpack.config.js",
        "watch-poll": "npm run watch -- --watch-poll",
        "hot": "cross-env NODE_ENV=development
node_modules/webpack-dev-server/bin/webpack-dev-server.js
--inline --hot --config=node_modules/laravel-
mix/setup/webpack.config.js",
        "prod": "npm run production",
        "production": "cross-env NODE_ENV=production
node_modules/webpack/bin/webpack.js --no-progress --hide-
modules --config=node_modules/laravel-
mix/setup/webpack.config.js"
    },
    "devDependencies": {
        "axios": "^0.17",
        "bootstrap-sass": "^3.3.7",
```

```
        "cross-env": "^5.1",
        "jquery": "^3.2",
        "laravel-mix": "^1.0",
        "lodash": "^4.17.4",
        "vue": "^2.5.7"
    }
}
```

phpunit.xml

```
<?xml version="1.0" encoding="UTF-8"?>
<phpunit backupGlobals="false"
         backupStaticAttributes="false"
         bootstrap="vendor/autoload.php"
         colors="true"
         convertErrorsToExceptions="true"
         convertNoticesToExceptions="true"
         convertWarningsToExceptions="true"
         processIsolation="false"
         stopOnFailure="false">
    <testsuites>
        <testsuite name="Feature">
            <directory
suffix="Test.php">./tests/Feature</directory>
        </testsuite>

        <testsuite name="Unit">
            <directory
suffix="Test.php">./tests/Unit</directory>
        </testsuite>
    </testsuites>
    <filter>
        <whitelist
processUncoveredFilesFromWhitelist="true">
            <directory suffix=".php">./app</directory>
        </whitelist>
    </filter>
    <php>
```

```
        <env name="APP_ENV" value="testing"/>
        <env name="CACHE_DRIVER" value="array"/>
        <env name="SESSION_DRIVER" value="array"/>
        <env name="QUEUE_DRIVER" value="sync"/>
    </php>
</phpunit>
```

APPENDIX B – Implementation act

Товарищество с ограниченной ответственностью «Astana IT University»

ASTANA IT UNIVERSITY

Жауапкершілігі шектеулі серіктестік «Astana IT University»

Қазақстан Республикасы, Нұр-Сұлтан қаласы, Мәңгілік ел даңғылы, С1 павильоны
Тел.: +7(717) 245 6292, +7(727) 327 5455
Email: info@astanait.edu.kz

Republic of Kazakhstan, Nur-Sultan, Mangilik el avenue, pavilion C1
Tel.: +7 (717) 245 6292, +7 (727) 327 5455
Email: info@astanait.edu.kz

Республика Казахстан, г. Нур-Султан, проспект Мәңгілік ел, павильон С1
Тел.: +7(717) 245 6292, +7(727) 327 5455
Email: info@astanait.edu.kz

02.12.2020 № 01/2301

АКТ
о внедрении результатов докторской диссертационной работы
Смайыл Асель Маралбайқызы

на тему: "Модели и методы интеллектуальной информационно-обучающей системы"

Результаты исследований диссертационной работы применяются в образовательной деятельности "Astana IT University" по следующим направлениям:

-проектирование и внедрение предлагаемой системы управления образовательного ресурса;

-разработка контента учебной практики для студентов 1 курса ОП "ИКТ".

При этом важно отметить, что практической ценностью исследования, выполняемого докторантом Смайыл А.М. является методика построения образовательного ресурса с помощью семантического анализа. Модели и методы формирования контента предназначены для пользователя, где учитываются особенности обучающего. Результаты исследований разработанной системы, показывающие повышение качества образования и эффективности за счет семантического представления ресурсов, также создает условия для цифровизации учебного процесса. Полученные результаты с возможной перспективой развития теоретической и прикладной базы автоматизации процесса обучения путем учёта индивидуальных особенностей каждого обучаемого студента. Использование полученных результатов позволяет: обеспечить наилучший способ обучения пользователей, эффективно использовать время на обучение по типу восприятия и индивидуализация траектории обучения для каждого пользователя.

Декан
ф-м.н., ассоц. профессор **М.Сергазиев**

Координатор ОП «ИКТ»
PhD, ассоц. профессор **Д.Едилхан**

APPENDIX C - Certificate of entering information into the state register of rights to objects protected by copyright

ҚАЗАҚСТАН РЕСПУБЛИКАСЫ

РЕСПУБЛИКА КАЗАХСТАН

СВИДЕТЕЛЬСТВО

О ВНЕСЕНИИ СВЕДЕНИЙ В ГОСУДАРСТВЕННЫЙ РЕЕСТР ПРАВ НА ОБЪЕКТЫ, ОХРАНЯЕМЫЕ АВТОРСКИМ ПРАВОМ

№ 19339 от «14» июля 2021 года

Фамилия, имя, отчество, (если оно указано в документе, удостоверяющем личность) автора (ов):
СМАЙЫЛ АСЕЛЬ МАРАЛБАЙКЫЗЫ

Вид объекта авторского права: **программа для ЭВМ**

Название объекта: **Интеллектуальная информационная система с использованием моделей и методов семантического представления данных**

Дата создания объекта: **03.04.2019**

Құжат тұлнұсқалығын http://www.kazpatent.kz/ru сайтының "Авторлық құқық" бөлімінде тексеруге болады https://copyright.kazpatent.kz

Подлинность документа возможно проверить на сайте kazpatent.kz в разделе «Авторское право» https://copyright.kazpatent.kz

Подписано ЭЦП

Осланов Е.К.