VISUALIZATION OF DATA ANALYSIS FOR EVALUATION OF THE LEVEL OF IT-SPECIALISTS COMPETENCIES' CREATION

Dr. Oksana GUSHCHINA
ORCID:0000-0003-2381-8537
Department of Applied Mathematics and Computer Science
Togliatti State University
Togliatti, RUSSIA

Dr. Oksana ANIKINA
ORCID: 0000-0002-0676-0372
Department of Applied Mathematics and Computer Science
Togliatti State University
Togliatti, RUSSIA

ABSTRACT

In the study the authors show how visualization and analysis of education data could be used in simulation and adjustment of education program curriculum for future evaluation of the formation levels the graduates competences as specified by educational standard. The authors suggest the scenario that demonstrate the possibility of automatic evaluation of graduate competences formation. This study is a viable teaching tool that can be used to simulation of the pedagogical activity scenario in order to eliminate all kinds of gaps in the training of the graduates and the evaluation of their professional competencies. The approach proposed in the study can be used to build a more result-oriented educational program by its redesigning to eliminate identified gaps in the consideration of important learning outcomes (the formation level of graduates' professional competencies) and evaluate the educational program using an updated visual model of evaluation activity. The relevance of this work is also based on its contribution to expanding the capabilities of multidimensional analysis and visualization of the results in relation to solving professional educational tasks.

INTRODUCTION

Information technology education (IT) is an area that constantly needs to be evaluated and transformed in order to align with the quickest turnaround rate in various market areas. The main factor that fundamentally changes the training method is a technology (Ukuev, 2018). Examples of various types of technologies used in education are mobile devices, systems for teleconferences, remote access systems, learning platforms and others with which students, lecturers, and the administration of the education institution interact. These technologies are used to influence and improve teaching and training, and also show techniques for using modern technology under real-world conditions. Interaction with these technologies generates large amounts of data, which can be divided into the following groups (Utyomov and Gorev, 2018):

- personal data;
- data on the interaction of students with electronic learning systems (electronic textbooks, online courses);
- data on the effectiveness of teaching materials;
- administrative (total-system) data;
- predicted data.

Nevertheless, educational systems are not fully prepared to handle with these data, and use these systems in order to continuously improve the quality and evaluate the level of formation graduates' competences. Particularly, education in IT training for students constantly needs to consider the high volume of changing knowledge and actual data in learning activity in order to meet the needs of qualified IT specialists that are capable of solving professional tasks.

All data are stored in various structured and unstructured formats. Some data are available, while other often require more effort to its processing. Moreover, the data sources are often fragmented and not available to those who need them. Even when data are available, decision makers often do not know about it and do not have the tools or necessary skills to use the data (Cech et al, 2018; Custer et al, 2018). At the same time if both teacher's knowledge and available data are used, students can get better results in the process of education (Blazar, Kraft, 2017).

Two analytical methods (educational data mining and training analytics) that can upgrade training systems and success level of educator/educational institution are used to improve the quality of educational services and various key performance indicators (Adejo and Connolly, 2017; Liñan and Pérez, 2015). In cases where data mining tends to focus on developing new tools for discovering patterns in data, training analytics focuses on applying tools and techniques at larger scales (Oyerinde et al, 2015).

Big data in higher education could allow developing an understanding of "student achievement and training approaches" and having positive influence on key areas such as actual student achievement in accordance with the curriculum (Vaitsis et al, 2016). Big data and analytics in higher education have recently been seen as holding great potential to promote actions concerning 'administrative decision-making and organizational resource allocation', early identification of at-risk students and interventions to prevent them from failing, the development of more effective instructional techniques, and transforming the traditional view of the curriculum into a network of relations, using educational data collected regularly from learning management systems, social networks, learning activities, and the curriculum itself (Masino and Niño-Zarazúa, 2016; Braganza et al, 2017). The curriculum, its content and the learning outcomes are part of the educational data on the basis of which analytics can be used to study quality and higher education improvement and the level of graduates competence evaluation (Agasisti and Bowers, 2017; Kumar and Vivekanandan, 2018; Kulasegaram and Rangachari, 2018).

The emergence of big data requires new prospects for data management and analytics, including advanced methods and tools for data visualization to support learning processes. The developing research area in the field of visual analytics has the benefit of combining data analysis and data manipulation, representation of information and knowledges for the understanding and recognition of visual patterns.

The purpose of the study is to show effectiveness of the software tools of the visualization of educational data analysis for adjustment of education program that is aimed at competencies' creation of students.

## JUSTIFICATION AND PROBLEM STATEMENT OF THE RESEARCH

Large amounts of educational data are captured and generated on a daily basis from different sources and in different formats in the higher educational ecosystem (Vaitsis et al, 2016; Murumba and Micheni, 2017). Educational data vary from data obtained from the use and interaction of students with Learning Management Systems (LMS) and platforms, to learning events and course information, including the curriculum, learning materials, exam results and other types of data related to administrative, educational and qualitative processes and procedures. In addition, continuously updated technologies and approaches in IT, which should be reflected in educational activities, develop a lot of data, and its nature is not static and is like a snapshot (Davari et al, 2019) of a long-term volatile network at the time of its aquisition.

In modern conditions, the development of information technologies requires graduate to comply with a high level of professional readiness. The advent of professional standards gives to the employer the opportunity to fix the requirements for graduates' knowledge and skills that define the specialist (Rehman, 2014). Requirements exist for theoretical knowledges and for practical skills of applicants. On the other hand, educational institutions develop required competencies for students through the Federal State Educational Standard, curriculums, and education plans. The result of joint activity of potential employers in the educational process is a specialist, an active person who is able to define and achieve the aim of one's professional activity.

Table 1 presents the average results of expert assessments obtained from 100 respondents that assessed the level of the impact of modern areas of information technology development on a 10-point scale, where 1 is a little impact and 10 is the strong impact. An expert assessment of the potential impact of the IT development on the formation of educational programs in the field of information technology was calculated on the basis of these indicators (Figure 1).

Table 1. Experts assessment results of modern IT influence on IT educational programs and students' competencies

IT research fields		IT degree	Competencies	Knowledge	Efficiencies	Skills
Artificial intelligence	d1	3,5	0,35	1,2	0,6	1,8
Bioinformatics	d2	2,45	0,65	1,05	0,3	0,4
Computational science	d3	1,65	0,85	0,15	0,2	0,4
Computer science	d4	2,45	0,445	0,9	0,3	0,4
Database engineering	d5	3,5	0,3	1,35	0,8	1,6
Graphics	d6	2,8	0,7	0,15	0,8	1,6
Human-computer interaction	d7	1,35	0,8	0,15	0,2	0,2
Information science	d8	2,75	0,6	0,15	0,8	1,8
Knowledge engineering	d9	3,25	0,35	1,2	0,8	1,8
Multimedia design	d10	2,45	0,65	1,35	0,6	1,6
Network engineering	d11	2,75	0,45	1,5	0,3	0,4
Performance analysis	d12	2,45	0,55	1,5	0,6	0,8
Software architecture	d13	1,4	0,85	0,15	0,2	0,4
Software engineering	d14	1,35	0,445	0,9	0,3	0,4
System administration	d15	1,55	0,3	1,35	0,8	1,6
System security and privacy	d16	3,15	0,7	0,15	0,8	1,6
Web service design	d17	2,45	0,8	0,15	0,2	0,2
Robotics and robot software	d18	3,5	0,6	0,15	0,8	1,8
AR and VR	d19	2,8	0,35	1,2	0,8	1,8
Cloud technologies	d20	1,35	0,65	1,35	0,6	1,6
Smart cities	d21	2,75	0,45	1,5	0,3	0,4
Data science	d22	3,25	0,55	1,5	0,6	0,8
Internet of things	d23	2,45	0,85	1,35	0,7	1,2
Big data	d24	2,75	0,35	1,8	0,5	1,4
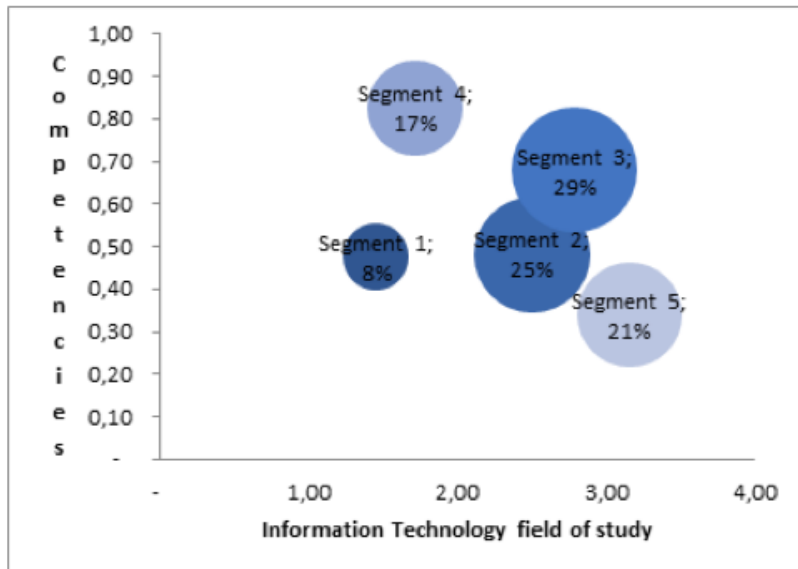
Figure 1. Influence cluster of modern IT on IT educational programs and students competencies formation

Figure 2. shows the dependence diagram of IT specialization, professional competencies and modern lines of IT research.
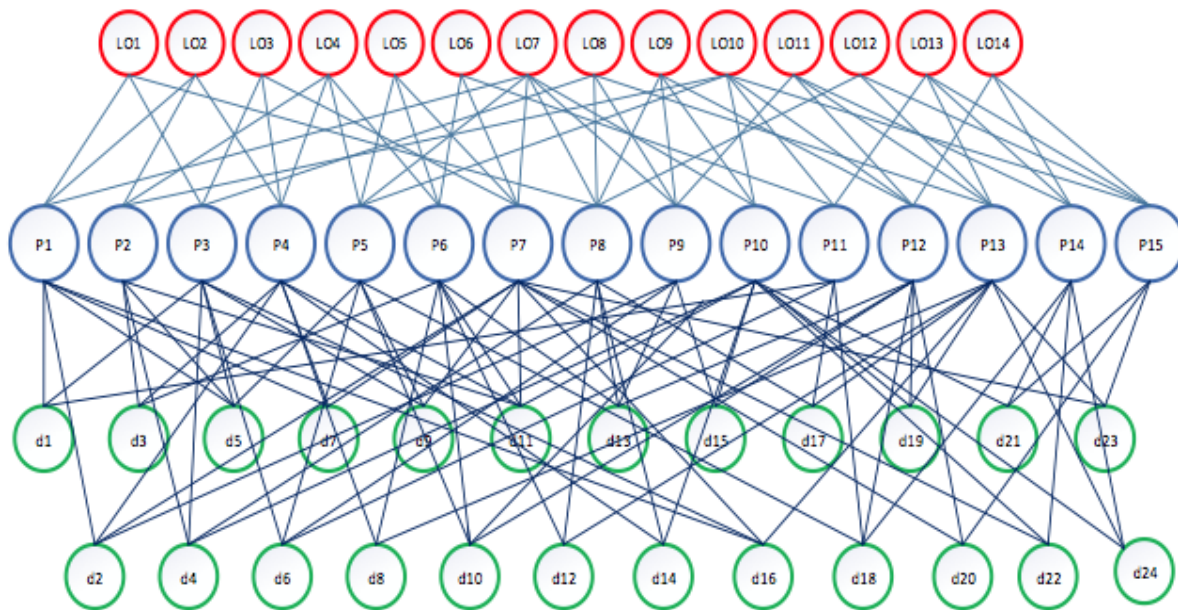
Figure 2. Dependence diagram of IT specialization, competencies and IT sphere

Tables 2-3 show key measures presented in Figure 2.

Table 2. IT specialization

IT specialization

Software programmer	P1
Software architect	P2
Test leader	P3
Database manager	P4
Software product manager	P5
Information Resources Specialist	P6
Information Technology Manager	P7
IT system technician	P8
Technology leader	P9
Product owner	P10
Systems analyst	P11
Information technology systems administrator	P12
System programmer	P13
Software/Web developer	P14
Systems integration specialist	P15

Table 3. Description of results of education for IT educational programs

Professional competencies	
The ability to collect, process, and interpret data of modern scientific studies necessary to form conclusions on the scientific studies	LO1
The ability to understand, upgrade and use of modern mathematical tool	LO2
The ability to develop and use of algorithmic and programming solutions in the area of system and applied software	LO3
The ability to apply modern information technology in designing, implementing, assessing and analyzing of the software efficiency for solving problems in different subject areas	LO4
The ability to use basic techniques and design automation facilities, implementation, investigation and quality assessment when creating a competitive software product and software systems, and methods and tools of automating associated with the maintenance, administration and modernization of software products and software systems	LO5
The ability to use the knowledge of computers trend with traditional (non-traditional) software design; modern system software; operating systems, operating and network shells, service programs; development trends of functions and design of problem-oriented software systems and complexes in professional activities	LO6
The ability to use the conceptual settings of the functional, logical, object-oriented and visual programming areas, methods, techniques and development tools of programs within these areas	LO7
The ability to use modern development methods and implementation of specific algorithms of mathematical models based on software languages and application software package of modeling	LO8
The ability to make a survey of organizations, identify information needs of users, formulate requirements for the information system	LO9
The ability to develop and adapt application software	LO10
The ability to design IP by type of software	LO11
The ability to compose a technical and economic feasibility study for design decisions and terms of reference for the design of an information system	LO12
The ability to test IP software components	LO14
The ability to maintain a database and support information application for solving applied problems	LO14

Significant estimates can be drawn on the basis of the presented visualization, for example, which ability is most affected by the educational program (specialty of IT training) and/or by the field of professional research (main trends in IT). You can also demonstrate the clusters of the educational program used to solve professional problems in the field of modern IT areas, and more.

Thus, large educational data is simulated by information visualization and data analysis (VA) methods and presented in visual interfaces with which human visual perception interacts to influence the process of analytical thinking (Khan and Khan, 2011; Cashman et al, 2018; Keim et al , 2008; Choi et al, 2019).

VA has the potential to support the process of manipulating and using big data by creating a whole picture of the data, uncovering the basic complex information to the greatest extent possible in order to have a positive impact on analytical thinking and decision making.

The most characteristic of the information visualization tasks and analysis of educational data are the following (Gorlushkina et al, 2015):

- monitoring and evaluation of the professional competencies formation;
- curriculum and course development;
- analysis and forecasting of students competitive ability in the labor market;
- diagnosis of the educational program quality level;
- assessment of the education quality for its compliance with standards.

The general objective of the study is to research new ways of analyzing and presenting the curriculum data of the educational program and the results of its development, aimed at assessing the level of graduates' professional competencies formation, using methods of visualization of data analysis (VA).

Specific objectives are: (1) signification of the various aspects affected to how the training is conducted by understanding the collected educational data from the educational program to determine and correctly apply the methods for analyzing and identifying important aspects in them; usage of VA for further analysis and visualization of the identified aspects and determining the value of VA methods applied to the curriculum data of the educational program; (2) construction of a mathematical model of graduate professional competencies evaluation to develop a visual evaluation method of the level of graduates' professional competencies formation for the implementation of the prototype automation software tools for the visualization process of information and analysis of educational data.

APPROACHES TO THE RESEARCH PROBLEM

Analysis of the Educational Program Curriculum Data and Its Visualization

An analysis of educational data was made to create a scientific basis and determine the scope of visualization within the curriculum (course) of the educational program. The analysis includes various learning activities (teaching methods), assessment methods (written and other types of exams), learning outcomes (LO1 – LO14) and the main results (knowledges (Z), efficiencies (U) and skills (V)).

Figure 3 illustrates how the learning outcomes (LO) and training methods (TM) of one course were simulated to visually represent a hidden network and relationship between them.
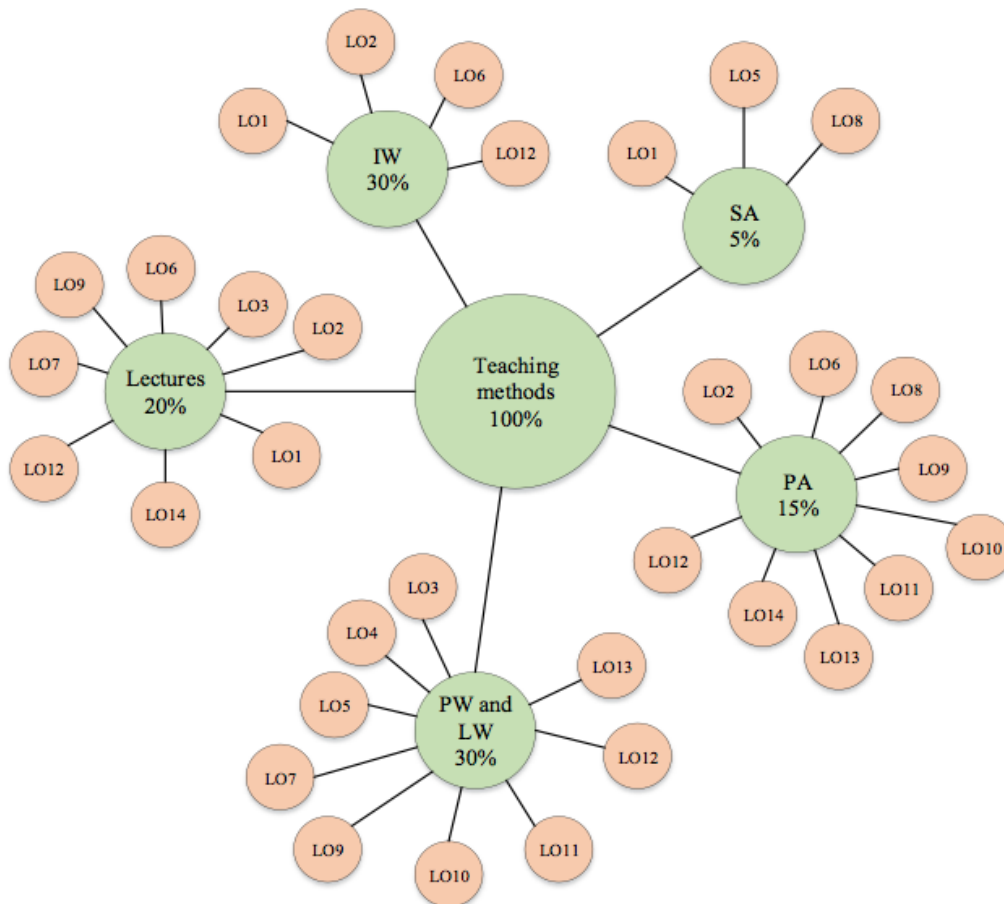
Figure 3. The relationship of the learning outcomes and teaching methods within the framework of the curriculum

Visualization of training methods (TM) (indicated in green) shows the use of each TM in the course. It is shown in what proportion a particular training method is involved in the IT graduate professional competencies formation for each TM. The diagram shows that the TM includes lectures, practical and laboratory work (PW and LW), self-directed learning (SL), as well as project (PA) and scientific (SA) activities. It can be said that the main contribution to the formation of the estimated results is made by practical and laboratory works, as well as self-directed learning. Project activity is an important part of pedagogical methods oriented to consolidate the material received from the learning process; it occupies 15%.

Thus, it is possible to review the structure, interconnection and correlation of learning methods with regard to the results, as well as to determine the share of participation of each training methods in their formation.

Figure 4 presents the modeling of the learning outcomes of one educational program course. The percentages in green circles represent the share of various types of final control according to the results of the educational program learning. It can be seen that the main type of the results examination is testing; an integral part of the final control for testing practical skills is the existence of term papers (TP) or course projects (CP), as well as the provision of reports on the laboratory (LW) and/or practical work (PW).

Testing results allow to assess the level of new knowledge and skills, and the fulfilment of course projects and practical works are the results of practical training. The final test is carried out in the form of computer testing and includes three main modules: the first module (M1) represents the theoretical aspect aimed at knowledge checks; the second module (M2) is the practical part that checks the level of theoretical material learning and the formation of skills; the third module (M3) is a control unit and is aimed at skills checks.
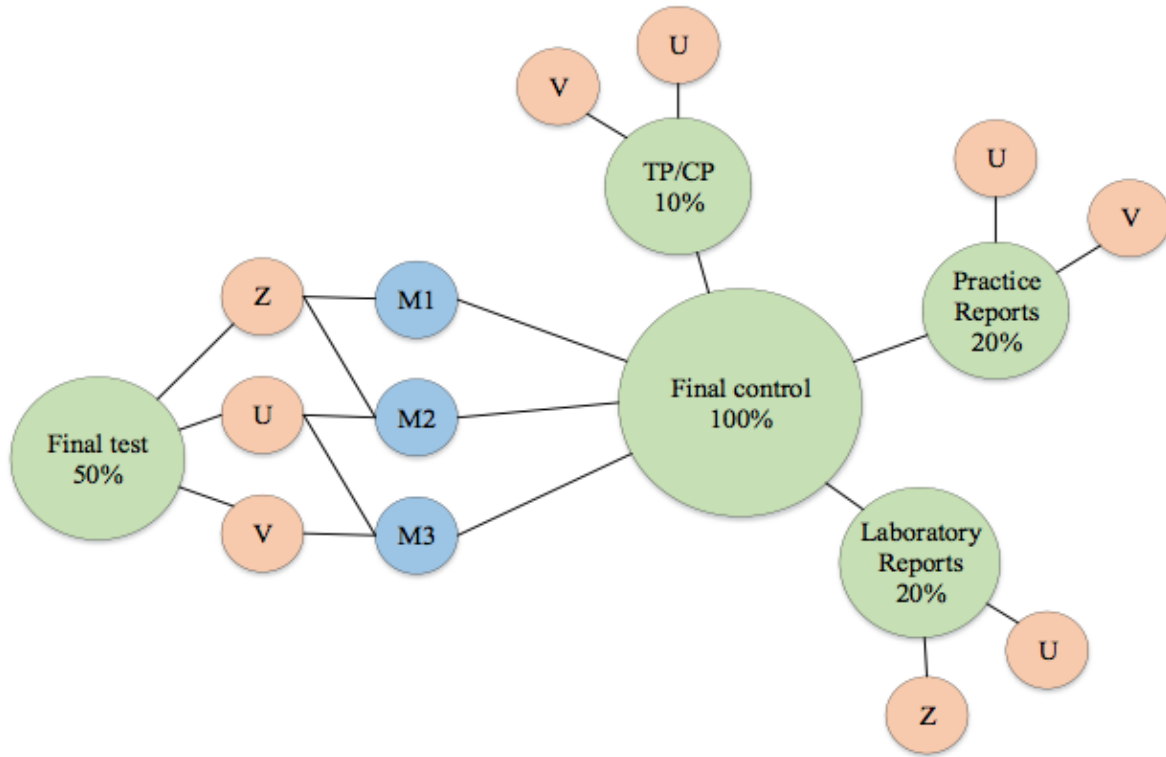
Figure 4. Final control and learning outcomes

This approach can be used to make a more exam-oriented educational program by redesigning it to eliminate identified gaps in the consideration of important LOs and evaluate the educational program using an updated visual model of evaluation activity.

Figure 5 shows the visualization of the entire educational program. Training methods are shown in green, LOs are shown in blue, and the main results aimed at determining the conformity of the chosen direction in the field of information technology and the educational program are shown in yellow. Split lines show how teaching methods influence on the formation of professional competencies (LO). The main results of the mastering the educational program are formed in the areas of IT research (di). The diagram shows the results of mastering the Applied Informatics curriculum, two study groups with an average of 20 people participated in the testing. The diagram (Figure 5) shows that the results of mastering the educational program and the level of students' competencies make part of the specialists in modern areas of information technology presented in Table 1. The success rate is shown in percentage, the extent of the generated results in IT areas (di) correspond to that or other IT professional qualifications (Pi).
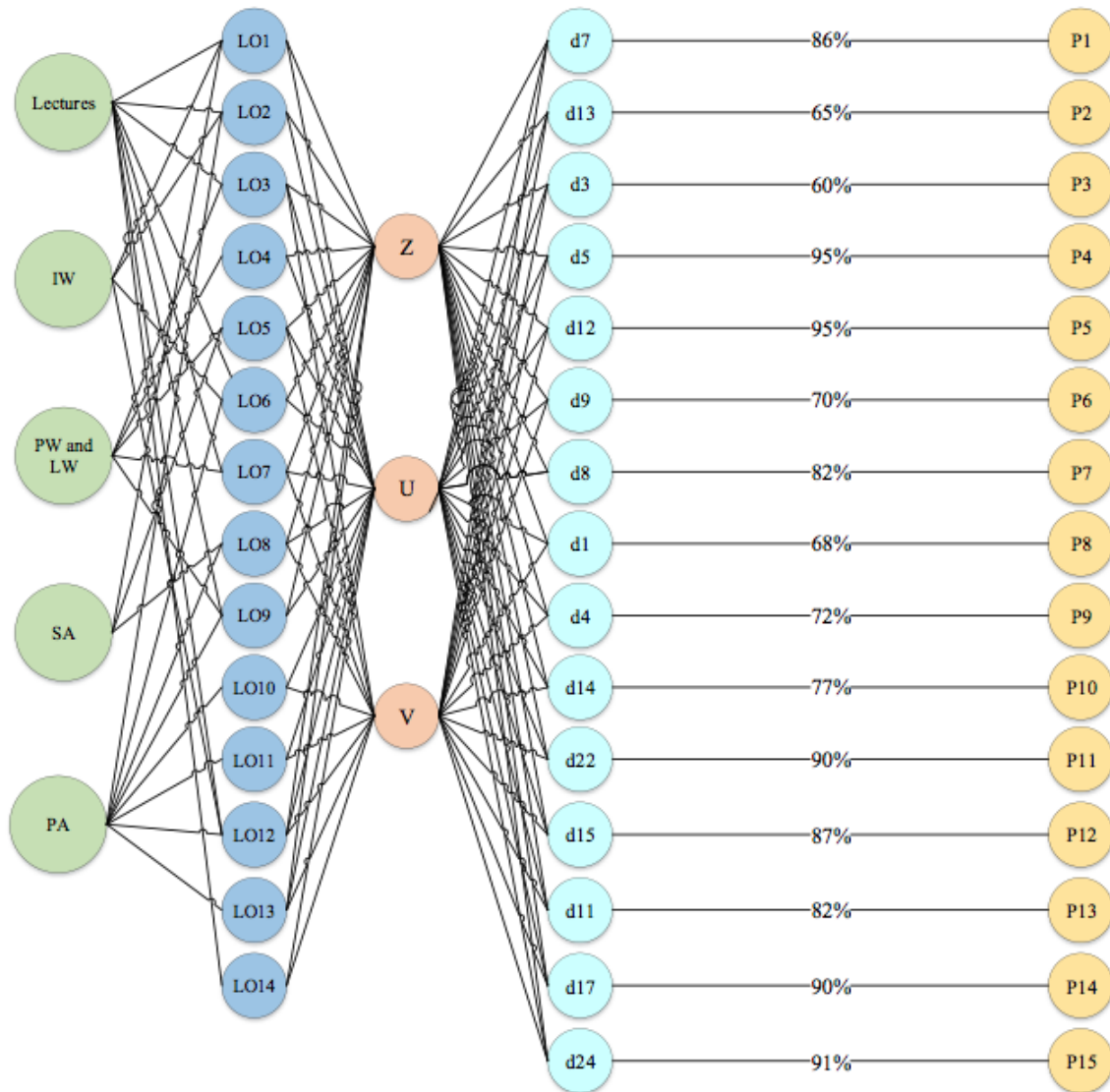
Figure 5. Structural alignment and educational program gap analysis

The realized visualization summarizes all the information from Figures 3 and 4, providing further more information about the educational program, so the entire educational program could be viewed as a whole and from different points of view.

In considering the model (Figure 5), it is possible to analyze educational data in different ways. One can see the relationship between the most specific and appreciated LOs, showing that the educational program aimed at knowledge, efficiencies and skills to determine the degree of their participation in the learning outcomes formation, as well as in identifying unformed learning outcomes and its gaps. The possibility that the structural alignment in the educational program can be tested as a synthesis of possible identified gaps and the use of learning events and LO in one place, presenting the educational program in the form of a structured system. The level of students' success in any of the learning outcomes can be associated not only with the importance of this learning outcome, but also with the degree to which it is taken into account in teaching methods. In this case, any gaps in the taught but unappreciated learning outcomes can be quickly identified.

Thus, at the stage of curriculum development, one can use mapping tools to identify actual gaps: determining teaching goals that cannot be solved correctly in the teaching or learning process; developing recommendations for determining new and motivating learning events to obtain the required learning outcomes.

In addition, with the help of analytics tools, one can further analyse a group of students and predict their needs such as performance, various conceptions of learning, technologies and group behavior. Several algorithms and prediction models that can develop the characteristics of a students' group and determine the degree of their professional competencies formation process data of this type. Visualization tools can be used to give alternative suggestions for developing proper types of activities for a particular group, as well as illustrate the effects of each option. In addition, visualization can clearly demonstrate the results of the knowledge formation, efficiencies and skills of students generated by training under the educational program, as well as show their level of competencies formation.

Analysis of Educational Data in The Construction of A Mathematical Model of Student Professional Competence Evaluation

The main information providers for the formation of graduates' competencies requirements are employers, employees and analysis of the workers functional responsibilities in the area specialist is training. Information fusion allows to form the requirements that the labor market applies to the competencies of graduates.

The problem of quantitative evaluation of the level of competencies and other students' qualification profile arises when developing an educational program for IT specialist with employers (Fomina et al, 2019; Leontyev et al, 2016; Zavadský et al, 2019). This is the relevance of the task of improving the mathematical tool for measuring and evaluating qualification characteristics, building models that allow doing the analysis of dynamics in student awareness level depending on the adoption of certain managerial solutions.

The formation of qualification units QuS can be realized in the process of learning the teaching units (modules), separate discipline or group of disciplines. At the same time, some competency-based qualification units, that fill the PC competence, are formed. In the process of learning the discipline Di, students get the knowledge Z_{Di}, efficiencies U_{Di} and skills (proficiency) V_{Di}. The complex of knowledge, efficiencies and skills formed during the study of the discipline influences on the final indicator of competence in the form of one or more competence-based units.

Mathematical problem statement is defined as follows: each element of the Qui set is assigned a numerical score that reflects the level of the qualification unit, then the generalized qualification model of a specialist can be described as follows:

$Qu^i = \{qu^1, qu^2, …, qu^S\}$

Where

Qu^i is the description of the qualification unit;

i is its number in the set;

S is the total number in the set.

Figure 6 shows the scheme for filling competence with qualification units in the process of studying the disciplines of the educational program.

The result of the evaluation is a qualification model of a specialist, which includes:

skill profile, as a set of qualification units;

set of numerical scores describing the level of employers requirements.

The qualification model of a specialist determines the set of professional competencies $PC = \{PC^{[i]}\}, (i = \overline{1:S})$,

, which the student should get while learning. Each professional competence $PC^{[i]} = \{Z^{[i]}, U^{[i]}, V^{[i]}\}$ consists of ordered sets of "knowledge"

$Z^{[i]} = \left(z_1^{[i]}, …, z_n^{[i]}\right)$, "efficiencies" , "proficiencies" $U^{[i]} = \left(u_1^{[i]}, …, u_m^{[i]}\right)$,

$V^{[i]} = \left(v_1^{[i]}, …, v_r^{[i]}\right)$, that indicate the structure of the studied disciplines $D = \langle D_j \rangle, (j = \overline{1:T})$. .

The qualification model of a specialist can be represented as a space of sets:

{PC, Z, U, V, D, T}

Where

PC is a set of professional competencies;

Z is a set of "knowledge" elements;

U is a set of "efficiencies " elements;

V is a set of " proficiencies " elements;

D is a set of disciplines identified by the educational program;

T is the set of time response characteristic of the evaluation process.
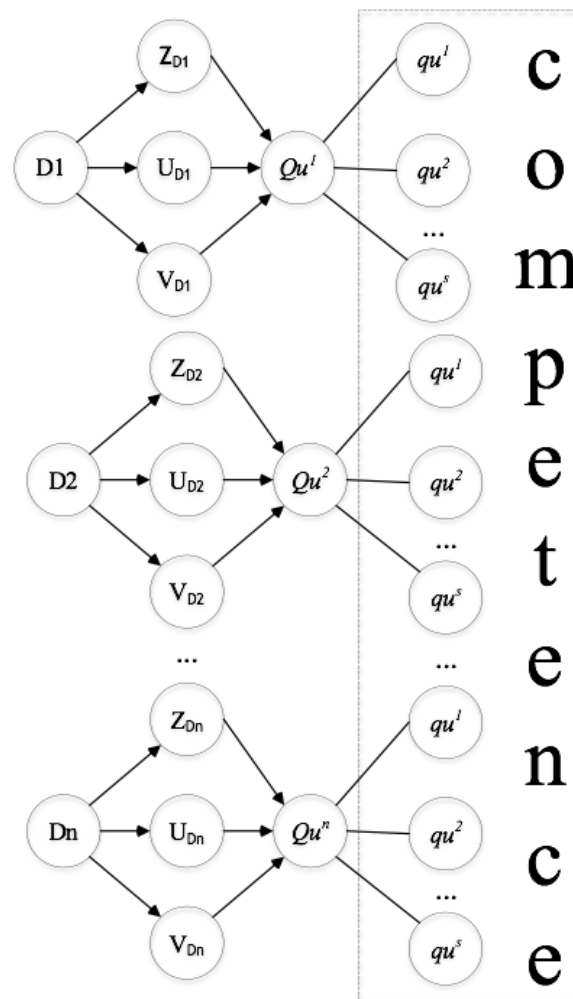
Figure 6. Competence Formation Scheme

An ordered set of all educational disciplines determines its competence based content. The proposed model (Figure 7) makes it possible to present the structure of the process of professional competencies and discipline formation in the education system through the interconnection of their components: "knowledge – efficiencies – proficiencies" (Gitman, 2014).
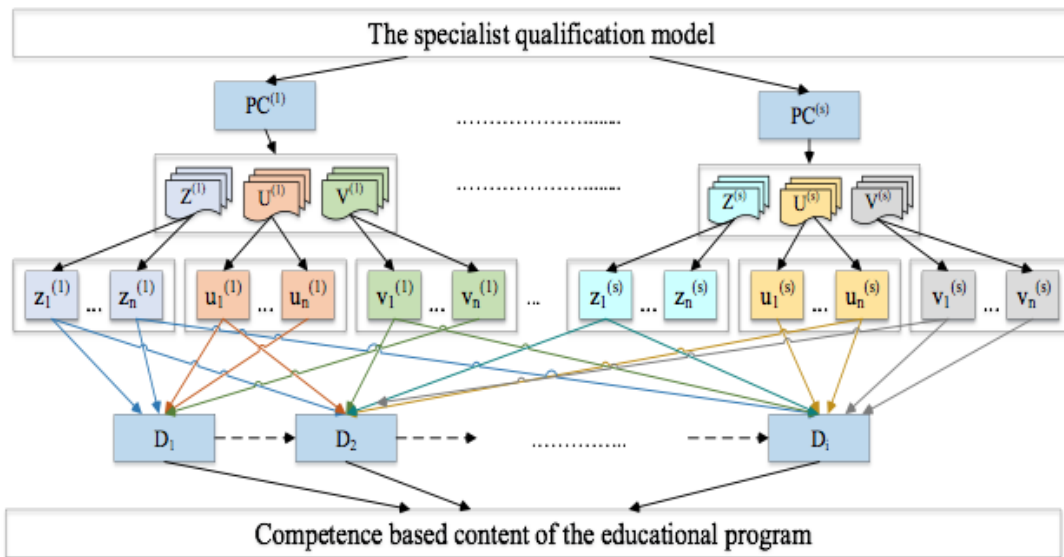
Figure 7. Competence based content of the educational program for the implementation of the specialist qualification model

An analysis of existing approaches to professional competencies evaluation made it possible to identify a generalized multicomponent mathematical of an estimation of professional competence (Borzykh et al, 2015), its result is determined by general grades for the student's academic achievements, personality evaluation and professional motivation for learning that is fully compliant with the requirements of the Federal State Educational Standard of Higher Education:

$$\overline{PC}_T^{[i]} = f\left(\overline{AA}_T^{[i]}, \overline{PQ}_T^{[i]}, \overline{PM}_T^{[i]}\right), (i = \overline{1:S}, T = const)$$

Where

$\overline{PC}_T^{[i]}$ **is** general assessment of the i professional competence in the T period;

$\overline{AA}_T^{[i]}$ **is** general assessment of academic achievements for the i professional competence in the T period;

$\overline{PQ}_T^{[i]}$ **is** general assessment of personality evaluation for the I professional competence in the T period;

$\overline{PM}_T^{[i]}$ **is** a general assessment of professional motivation for the i professional competence in the T period.

Thus, the task of the level of formation of student professional competencies evaluation is reduced to a phased presentation of the results of educational and other activities.

In the educational system, in the conditions of a competence based approach, the results of students' academic achievements should be presented in the form of formed competencies (general, professional, etc.) (Haddouchane et al, 2017; Muñoz and Araya, 2017), that determine the qualification model of a specialist, should be decomposed into the components "know – be able – own". Likewise, the results of disciplines development as well as competencies should be presented in the form of formed knowledge, efficiencies, proficiencies (Table 4).

Table 4. The result of the discipline development

Discipline	Components	Grades	Allied disciplines	Coefficient of the components significance
	knowledge	$z_{j_1}^{[i]}$	D2, D7, D11	$\alpha_1^{[i]}$
		...	...	...
		$z_{j_\alpha}^{[i]}$	D4, D8, D19	$\alpha_a^{[i]}$
Dj	efficiencies	$u_{j_1}^{[i]}$	D3, D5, D8, D21	$\beta_1^{[i]}$
		...	...	...
		$u_{j_\beta}^{[i]}$	D4, D11, D16, D21	$\beta_b^{[i]}$
	proficiencies/skills	$v_{j_1}^{[i]}$	D4, D18	$\gamma_1^{[i]}$
		...	...	...
		$v_{j_\gamma}^{[i]}$	D11, D15, D24	$\gamma_c^{[i]}$

Since the formed professional competencies are a system of knowledge, efficiencies and skills to apply them in practice in professional activities (Kadralinova, 2015; Leontyev et al, 2016), and "knowledge – efficiencies – proficiencies" are defined by some set of disciplines, in order to measure competences, it is necessary to determine the quantitative evaluation of knowledge, efficiencies, proficiences, taking into account the interconnections of disciplines (Figure 7).

Evaluation of student learning activities is a convolution of private evaluations of the i professional competence, that are studied and measured in different disciplines, they form the measured i professional competence:

$$
\begin{cases}
\overline{AA}_T^{[i]} = \sum_{j=1}^{\tau} \overline{AA}_{D_j}^{[i]}, (j = \overline{1:T}, i = \overline{1:S}), \\
\overline{AA}_{D_j}^{[i]} = \sum_{j=1}^{T} \sum_{k=1}^{a} \alpha_k^{[i]} \bar{z}_{j_k}^{[i]} + \sum_{j=1}^{T} \sum_{k=1}^{b} \beta_k^{[i]} \bar{u}_{j_k}^{[i]} + \sum_{j=1}^{T} \sum_{k=1}^{c} \gamma_k^{[i]} \bar{v}_{j_k}^{[i]}, \\
\alpha_k^{[i]} = \frac{q(z)_k^{[i]}}{T}, \beta_k^{[i]} = \frac{q(u)_k^{[i]}}{T}, \gamma_k^{[i]} = \frac{q(v)_k^{[i]}}{T},
\end{cases}
$$

Where

$(\overline{AA}_{D_j}^{[i]}$ (is assessment of the collection of components of "knowledge – efficiencies – proficiencies" by j discipline participating in the formation of the i professional competence;

$\alpha_k^{[i]}$, $\beta_k^{[i]}$, $\gamma_k^{[i]}$ is significance coefficients of the "knowledge – efficiencies – proficiencies" components of the j discipline participating in the formation of the i professional competence;

$\bar{z}_{j_k}^{[i]}$, $\bar{u}_{j_k}^{[i]}$, $\bar{v}_{j_k}^{[i]}$ is assessment of the appropriate "knowledge – efficiencies – proficiencies" components of the j discipline participating in the formation of the i professional competence;

$q(z)_k^{[i]}$, $q(u)_k^{[i]}$, $q(v)_k^{[i]}$ is the number of disciplines involved in the formation of appropriate "knowledge – efficiencies – proficiencies" components of i professional competence;

T is the number of all disciplines of certain educational programs;

S is the number of all professional competencies in one exact course;

a, b, c are the number of evaluation components of the j discipline;

⊠ is the number of disciplines that form the i professional competence.

Thus, evaluation of professional competencies is defined as the result of academic achievements and is a convolution of private evaluations ("knowledge", "efficiencies", "proficiencies") in reference to all educational disciplines that form considered competencies.

Software Implementation of Mathematical Model of Evaluation of Graduate Professional Competences Formation

There are both separate powerful analytical applications integrated with database management and LMS systems and tools that work directly in the Microsoft Office environment to solve the problem of evaluation of the levels of formation of professional competences. Programs are able to provide users with time-sensitive information that is necessary for effective decision-making on the choice of further directions for the development of an educational establishment development. The question remains relevant for educators to choose from existing data analysis tools and their adaptation for data processing for widespread implementation in the process of learning monitoring and managing the tools that will be available and easy to study and use.

An implemented prototype of software tools in the programming language R was proposed in this study to solve the problem of evaluation of the levels of formation of professional competences in the qualification model of a specialist. The choice of the prototype is due to the following reasons (Kabakov):

- It is a free, functional, open source programming language;
- R programming environment is cross-platform and the applications created in it can be used without modification in the operating systems Windows, Unix, Linux;
- the R language was developed with the focus on effective statistical processing of data bulk;
- the R language includes several powerful graphics systems and is currently one of the best multidimensional data visualization environments.

The lattice package was used to implement the required functions of displaying information in software tools for graphics output. The lattice package is a powerful and elegant high-level data visualization system with a focus on multidimensional data and aimed at developing Trellis graphics for R language to meet typical graphic needs with minimum setting.

The developed software tool provides an opportunity to analyze the formation level of competencies both taking together study groups within the field of study and the corresponding curriculum and within the specified group and the particular student. An expert assessment of students' professional competence formation was used to process large data samples and visualize the results of calculations

Figure 8 presents diagram that shows that all data processing work is performed automatically in a software script.
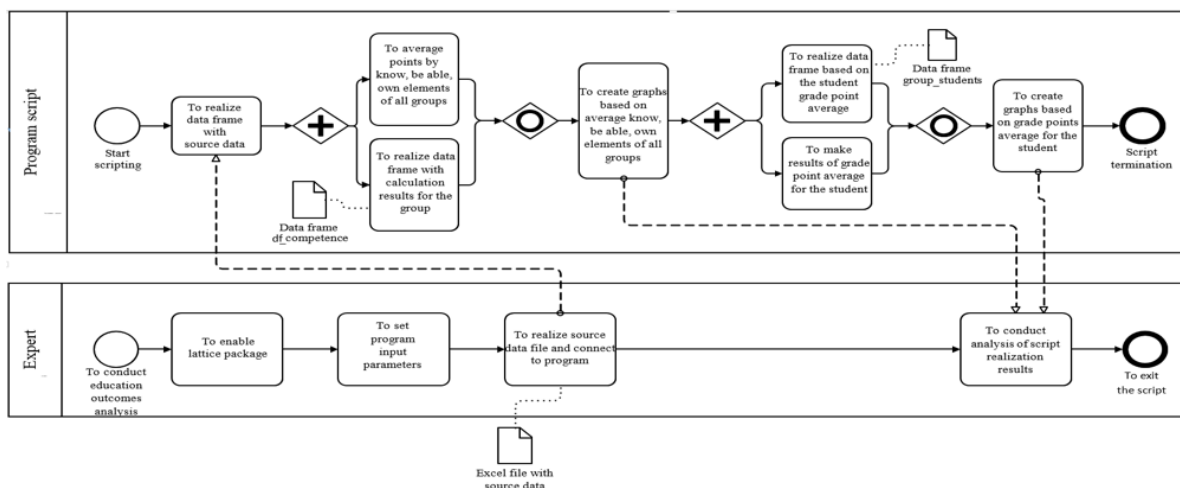

Figure 8. The procedure of program script on competence formation analysis

267

The implementation of software tools was based on an algorithm that includes the following steps:

(1) Preparation and loading of source data.

(2) Computation of the average grade of tested students on the elements «know», «be_able», «own»

(3) Analysis of drilling results and their visualization.

Formation of The Initial Sample for The Analysis of Educational Data

The file, that was generated in the Excel, was used as the source data. This file contains the information about the test results of 10 study groups with 11 courses of the current curriculum of students studying in IT faculties. During testing, students received final grades on a scale from 1-5 for each course ranked by three elements: «know», «be_able», «own».

In this study students are made anonymous for the most objective assessment. Figure 9 shows the first and the last five lines of the Source Data.xlsx file content.

	A	B	C	D	E	F	G	H	I
1	discipline	part	competence	group	student	know	be_able	own	num
2	IS-1	0.4	PK-4	Gr.01	Student_001	4	4	4	1
3	IS-1	0.4	PK-4	Gr.01	Student_002	4	4	4	2
4	IS-1	0.4	PK-4	Gr.01	Student_003	5	4	3	3
5	IS-1	0.4	PK-4	Gr.01	Student_004	4	4	4	4
6	IS-1	0.4	PK-4	Gr.01	Student_005	4	3	4	5
...									
2755	Modern database technology and in	PK-3		Gr.10	Student_193	4	4	3	13
2756	Modern database technology and in	PK-3		Gr.10	Student_194	5	3	4	14
2757	Modern database technology and in	PK-3		Gr.10	Student_195	5	5	5	15
2758	Modern database technology and in	PK-3		Gr.10	Student_196	5	5	4	16
2759	Modern database technology and in	PK-3		Gr.10	Student_197	4	4	4	17

Figure 9. Part of Source data.xlsx file (the first and the last five lines)

Data frame with the source data intended to data processing is updated at the beginning of the script in the global environment R. The df_competence data frame includes the name of the discipline, the name of the competence, the number of the student group, the symbolic name of the student, the values for the elements "know", "be able", "own", the serial number of the student, as well as the average for all of the elements. Figure 10 shows the first and the last five lines of the df_competence data frame.

	discipline	competence	group	student	know	be_able	own	num	mean
1	IS-1	PK-4	Gr.01	Student_001	4	4	4	1	4.000
2	IS-1	PK-4	Gr.01	Student_002	4	4	4	2	4.000
3	IS-1	PK-4	Gr.01	Student_003	5	4	3	3	4.000
4	IS-1	PK-4	Gr.01	Student_004	4	4	4	4	4.000
5	IS-1	PK-4	Gr.01	Student_005	4	3	4	5	3.667
...									
2754	Modern database technology and inform...	PK-3	Gr.10	Student_193	4	4	3	13	3.667
2755	Modern database technology and inform...	PK-3	Gr.10	Student_194	5	3	4	14	4.000
2756	Modern database technology and inform...	PK-3	Gr.10	Student_195	5	5	5	15	5.000
2757	Modern database technology and inform...	PK-3	Gr.10	Student_196	5	5	4	16	4.667
2758	Modern database technology and inform...	PK-3	Gr.10	Student_197	4	4	4	17	4.000

Figure 10. Part of df_competence data frame (the first and the last five lines)

Thus, a file was generated, that contains the results of students testing for their subsequent processing. Trough the software tools work, this file was transformed into a one that can work through the programming language R.

Data Processing with The Use of Software Tools

The calculation of the students grade point average within the formed competencies for the elements "know", "be able", "own" is performed in the following commands (Listing 1).

Listing 1

```
# Computation of the average on the elements «know», «be_able», «own» in each group
students.cnt <- length(unique(df_competence[,"student"]))
num <- df_competence[1:students.cnt,"num"]
students <- aggregate(cbind(know, be_able, own, mean) ~ group + student,
                data = df_competence, mean)
group <- students$group[students$student==student]
group.students <- students[students$group==group,]
rownames(group.students) <- 1:nrow(group.students)
```

The calculation results (Listing 1) are represented in graphs that allows to evaluate visually students' competencies formation.

A data frame df_student is created to receive the information about the student through the software script (Listing 2). Graphics rendition of the scripting result is shown in Figure 11.

Listing 2

```
# Data analysis of competencies' formation of the student
df.student <- df_competence[df_competence$student == student,c(1:3, 5:9)]
group <- as.character(df.student$group[1])
num <- df.student$num[1]
caption <- paste(group,": ", num, ". ", student, sep = "")
cat("\n", caption, "\n", sep="")
cat("----------------------\n")
df.student <- df.student[,c("discipline", "competence", "know", "be_able", "own", "mean")]
print(df.student)
cat("\nKnow:")
cat("\n-------")
print(table(df.student$discipline, df.student$know))
cat("\nbe_able:")
cat("\n-------")
print(table(df.student$discipline, df.student$be_able))
cat("\nown:")
cat("\n-----------")
print(table(df.student$discipline, df.student$own))
```

	discipline	competence	know	be_able	own	mean
187	IS-1	PK-4	3	4	5	4.000
384	IS-2	PK-4	4	3	4	3.667
581	POIS-1	PK-6	3	3	4	3.333
778	POIS-2	PK-6	5	4	4	4.333
975	Computer graphics	OPK-3	4	5	5	4.667
1172	Computer graphics	PK-2	4	3	3	3.333
1369	Application programming	OPK-4	4	5	4	4.333
1566	Application programming	PK-8	3	4	3	3.333
1763	Web programming	PK-7	4	5	4	4.333
1960	Web programming	PK-8	4	4	5	4.333
2157	OOP-1	PK-6	5	4	4	4.333
2354	OOP-2	PK-6	3	4	4	3.667
2551	Design workshop	PK-8	4	4	5	4.333
2748	Modern database technology and informatio...	PK-3	3	3	4	3.333

Figure 11. The contents of the df_student data frame

The presented chunks of code show the process of the educational data processing for further analysis through the programming language R.

The visualization of the results of educational data analysis with the use of software tools

The results of knowledge, efficiencies and skills formation of students within professional competencies for 10 test groups are shown in Figure 12. The graphs depict the level of students professional competencies formation.
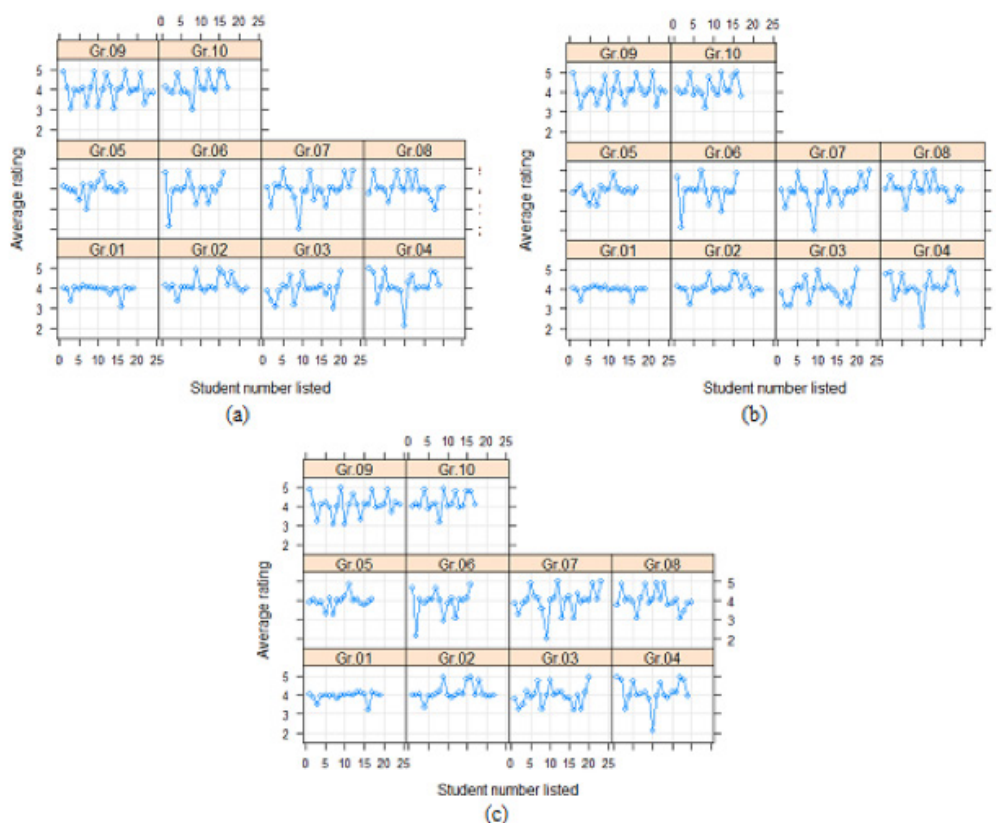
Figure 12. The average result of knowledge (a), efficiencies (b) and skills (c) formation for 10 test groups

270

So, for example, it can be seen that the average level of all students knowledge in group 01 is mostly within 4 points, i.e., competencies formation on the element of "knowledge" is at one level in this group. In groups 04 and 07 the average level of knowledge formation varies from 2 to 5 points, which shows different level of theoretical material learning within the group.

In the following graph (Figure 13): the average for the average results of the elements "know", "be able", "own".
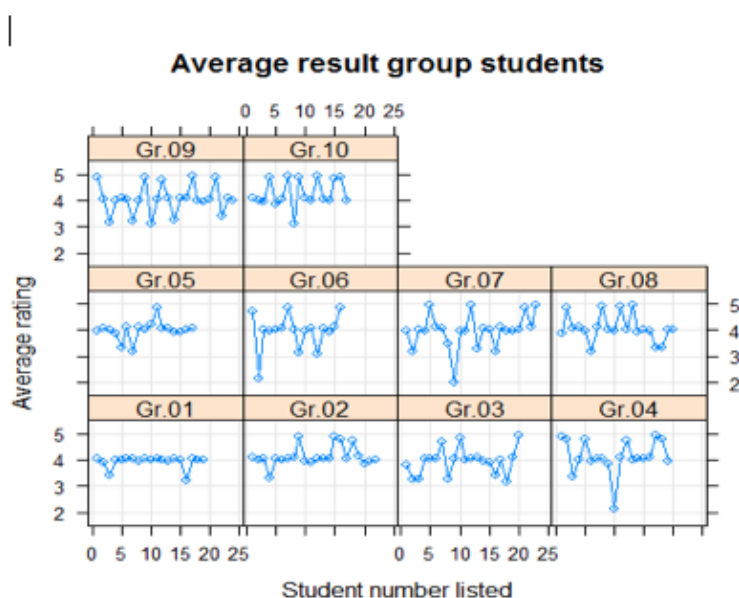
Figure 13. Average points for all competence elements of 10 test groups

Overall, the graphs allow to identify the weakest and the most powerful groups of students in professional competencies mastering. Thus, it can be determined that the most "strong" group after analyzing academic performance charts is 10th of the 10 tested groups. The main students contingent of the selected group has an grade point average of 4 or 5 for all "know", "be able", "own" elements, which show a high level of professional competencies formation and the disciplines mastering at a sufficiently high level. A similar conclusion can be on the 02th group.

At the same time, it is clear that the level of competence formation in 01th group is stably average. Average points for "know", "be able", "own" elements within the group is 4 points. The results of other groups range from 3 to 5 points, in some cases 2 points were fixed.

One can determine the level of particular student competencies formation by analyzing all tested groups and receive the display of information about the average points for "know", "be able", "own" elements in all disciplines of the current curriculum involved in the professional competencies formation (Figure 14).

```
> print(table(df.student$discipline, df.student$know))
```

	3	4	5
Application programming	1	1	0
Computer graphics	0	2	0
Design workshop	0	1	0
IS-1	1	0	0
IS-2	0	1	0
Modern database technology and information analysis	1	0	0
OOP-1	0	0	1
OOP-2	1	0	0
POIS-1	1	0	0
POIS-2	0	0	1
web programming	0	2	0

(a)

```
> print(table(df.student$discipline, df.student$be_able))
```

	3	4	5
Application programming	0	1	1
Computer graphics	1	0	1
Design workshop	0	1	0
IS-1	0	1	0
IS-2	1	0	0
Modern database technology and information analysis	1	0	0
OOP-1	0	1	0
OOP-2	0	1	0
POIS-1	1	0	0
POIS-2	0	1	0
web programming	0	1	1

(b)

```
> print(table(df.student$discipline, df.student$own))
```

	3	4	5
Application programming	1	1	0
Computer graphics	1	0	1
Design workshop	0	0	1
IS-1	0	0	1
IS-2	0	1	0
Modern database technology and information analysis	0	1	0
OOP-1	0	1	0
OOP-2	0	1	0
POIS-1	0	1	0
POIS-2	0	1	0
web programming	0	1	1

(c)

Figure 14. The average points of knowledge (a), efficiencies (b) and proficiencies (c) in all disciplines for Student_187

Figure 15 presents the average grade points of the student in all disciplines involved in the professional competencies formation.
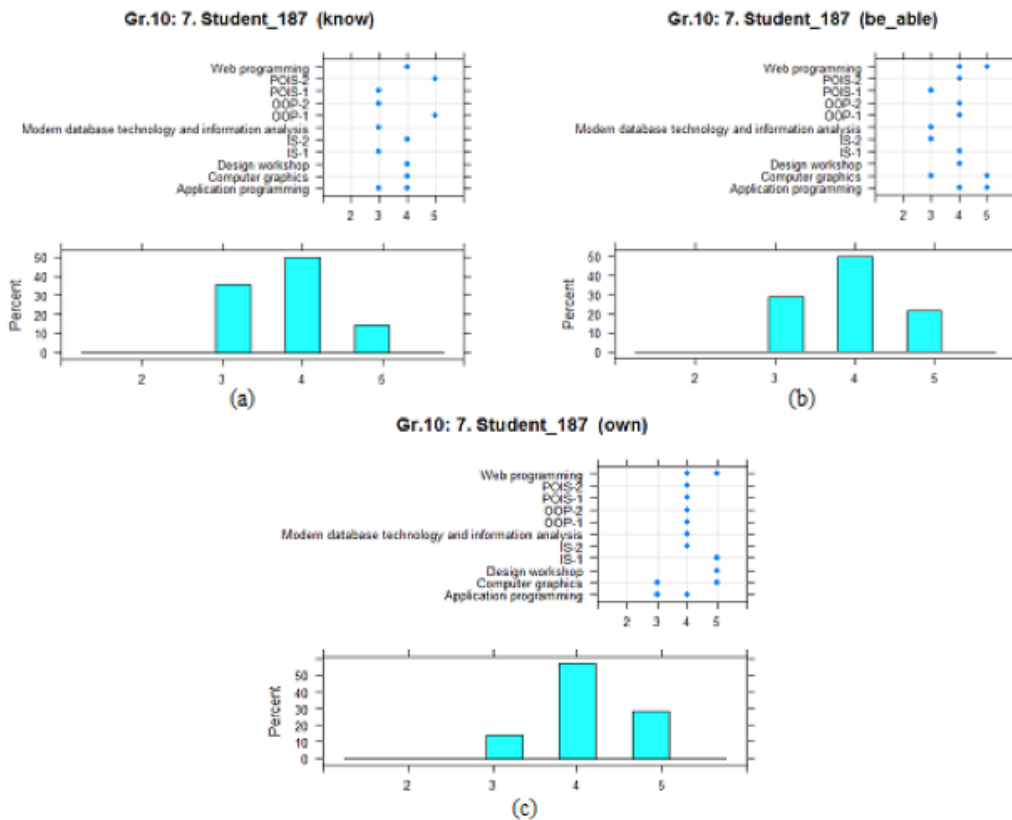
Figure 15. The average grade points of knowledge (a), efficiencies (b) and proficiencies (c) in all disciplines for Student_187

272

The presented graphs show the results of the student. It can be seen that in all disciplines of the curriculum there are mainly 5 points, in addition, there are no results below 4 points. These results demonstrate high grades of student Student_187 and show the the curriculum disciplines mastering mainly from 90 to 100%, which shows sufficiently high level of professional competence formation. The graphs visualize the results of the entire learning process and allow to conclude about the level of the student professional training based on a competence based approach.

Therefore, it is possible to analyze final grades of any student of the given sample. To process a new selection, one should enable at the beginning of the program script prepared for processing .csv file or a MySQL database.

The proposed script can be used for processing large amounts of educational data, and the powerful graphing capabilities of the lattice package allow to visualize the calculations results. Presentation of analysis results in the form of diagrams and graphs makes it possible to analyze the received information in visual mode, for example, to identify students status in the academic ranking within the group.

Developed application of academic performance evaluation can be used in the educational process to analyze students professional competence formation. It will allow to make timely adjustments to the organization of the educational process, plan the learning process in a specific discipline.

CONCLUSION

This study shows the influence of the results of the analysis and visualization of educational data on the process of the curriculum improvement and their reflection on the specialist qualification model formation organized on the basis of the competence based approach to graduate training in accordance with the educational standards requirements. The proposed approach can be used to develop more result-oriented educational program by redesigning it with consequent elimination of identified gaps in the consideration of important learning outcomes and educational program evaluation using an updated visual model of evaluation activity.

Different types of models providing the opportunity to observe and analyze the entire educational program from different points of view and in general were constructed with the usage of the visualization methods of educational data (Figures 1-5): to define the relationship between the most focused and evaluated learning outcomes by showing the educational program trend to knowledge, efficiencies and skills to determine the degree of their participation in the learning outcomes formation; to identify unformed learning outcomes and gaps. The proposed visual script model of the curriculum data analysis shows the ability to define presence or absence of lacks and contradictions in the existing curriculum for executive decision-making aimed at amendment and learning outcomes formation in accordance with educational standards.

Visualization of the need to redesign the IT specialist educational training program identified the demand for a quantification of the level of students' competencies used in the basis of the mathematical formulation of task of the specialist's qualification model organization, the model includes a set of qualification units and results of professional competencies formation evaluaton through the interconnection of their components: "knowledge – efficiencies – proficiencies".

A prototype of software tools in the programming language R is presented. The prototype allows to demonstrate the dynamics of curriculum disciplines mastering and the educational results of a student and a group on the basis of the mathematical formulation of task of the specialist's qualification model organization. This allows to see the results of the educational program mastering by students, to draw a conclusion about the level of graduates' professional competencies formation and to evaluate the efficiency of the program. The developed application for the students' professional competence formation based on the ranking of components assessments: "knowledge – efficiencies – proficiencies" helps to identify soft spots in the educational program for future improvement of the educational process organization and its planning.

The proposed script for the visualization automation and analysis of educational data will provide a better approach to the educational programs formation in the field of training of highly qualified specialists with the requierd rate of professional competences.

Thus, the visualization of educational data analysis is focused on the identification of gaps and inconsistencies in data, giving the opportunity to verify the conformity of learning activity with educational standards. The visualization of educational data analysis allows to easy acquire the structure of the data; to determine the reasons for their use and the importance of the obtained results. The visualization of educational data analysis helps to manage the divergences and structural defects, that were identified as a result of gap analysis, presence of constructive alignment in the data.

The results of the study can be used to model the script of teaching activities in order to eliminate all kinds of gaps in the preparation of graduates and their professional competencies evaluation. The proposed approach can be used to build more result-oriented educational program by redesigning it to eliminate identified gaps in the consideration of important learning outcomes (the level of graduates' professional competencies formation) and evaluate the educational program using an updated visual model of evaluation activity.

BIODATA and CONTACT ADDRESSES of AUTHORS

Oksana GUSHCHINA is a Head of the Department of Applied Mathematics and Computer Science, Institute of Mathematics, Physics and Information Technologies, Candidate of Pedagogic Sciences (2003), Associate Professor (2009). Her research interests are technologies of data science, e-learning and distance learning, AR and VR experience. She is the writer of some undergraduate programs in the field of Computer Science. Also, she is the author of more than 60 academic papers and 15 study guides. She has five academic papers were published in international academic journals.

Oksana GUSHCHINA
Department of Applied Mathematics and Computer Science
Institute of Mathematics, Physics and Information Technologies
Address: Togliatti State University 445020, 14, Belorusskaya St., Togliatti, Russia
Phone: +7 8482 546473
E-mail: ok_mih@mail.ru

Oksana ANIKINA is an Associate Professor at the Department of Applied Mathematics and Computer Science, Institute of Mathematics, Physics and Information Technologies, Candidate of Engineering Sciences (2012), Associate Professor (2019). Her academic interests include the analysis of educational data, distance technologies in the educational process, modern information technologies. She has more than 40 scientific papers have been published, including 2 monographs. She published two articles in international scientific journals and presented at international conferences.

Oksana ANIKINA
Department of Applied Mathematics and Computer Science
Institute of Mathematics, Physics and Information Technologies
Address: Togliatti State University 445020, 14, Belorusskaya St., Togliatti, Russia
Phone: +7 9277 862365
E-mail: blue-waterfall@yandex.ru

REFERENCES

Adejo, O., & Connolly, T. (2017). Learning Analytics in Higher Education Development: A Roadmap. *Journal of Education and Practice,* 8(15), 156-163.

Agasisti, T., & Bowers, A. J. (2017). Data Analytics and Decision-Making in Education: Towards the Educational Data Scientist as a Key Actor in Schools and Higher Education Institutions. *Handbook of Contemporary Education Economics*, 184-210. https://doi.org/10.7916/D8PR95T2

Blazar, D., & Kraft, M. A. (2017). Teacher and teaching effects on students' attitudes and behaviors. *Educational evaluation and policy analysis*, 39(1), 146-170. https://doi.org/10.3102/0162373716670260

Borzykh, V.E., Shalkina, T.N., & Nikolaeva, D.R. (2015). Method of mathematical modeling of graduates professional competencies evaluation process. *Modern problems of science and education*, 1-1, 106.

Braganza, A., Brooks, L., Nepelski, D., Ali, M., & Moro, R. (2017). Resource management in big data initiatives: Processes and dynamic capabilities. *Journal of Business Research*, 70, 328-337. https://doi.org/10.1016/j.jbusres.2016.08.006

Breiter, A., & Light, D. (2006). Data for School Improvement: Factors for designing effective information systems to support decision-making in schools. *Educational Technology & Society*, 9(3), 206-217. https://www.ifib-consult.de/publikationsdateien/Breiter_Light-2006.pdf

Cashman, D., Humayoun, S. R., Heimerl, F., Park, K., Das, S., Thompson, J., Saket, B., Mosca, A., Stasko, J., Endert, A., Gleicher, M., & Chang, R. (2019) A User-based Visual Analytics Workflow for Exploratory Model Analysis. *arXiv:1809.10782v3 [cs.HC].* https://arxiv.org/pdf/1809.10782.pdf

Cech, T., Spaulding, T. & Cazier, J. (2018). Data competence maturity: developing data-driven decision making. *Journal of Research in Innovative Teaching & Learning*, 11(2), 139-158.

Choi, I. K., Childers, T., Raveendranath, N. K., Mishra, S., Harris, K., & Reda, K. (2019). Concept-Driven Visual Analytics: an Exploratory Study of Model-and Hypothesis-Based Reasoning with Visualizations. *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, ACM, 68, 1-14. https://doi.org/10.1145/3290605.3300298.

Custer, S., King, E. M., Atinc, T. M., Read, L., & Sethi, T. (2018). Toward Data-Driven Education Systems: Insights into Using Information to Measure Results and Manage Change. *Center for Universal Education at The Brookings Institution*. https://www.brookings.edu/wp-content/uploads/2018/02/toward-data-driven-education-systems.pdf

Davari, M., Noursalehi, P. & Keramati, A. (2019). Data mining approach to professional education market segmentation: a case study. *Journal of Marketing for Higher Education*, 29(2), 45-66. https://doi.org/10.1080/08841241.2018.1545724

The Federal State Educational Standards of Higher Education (FSES HE), Retrieved 08 April, 2020 from http://fgosvo.ru/fgosvo/151/150/24

Fomina, L. V., & Shmeleva, Zh. N. (2019). Practical experience of training specialists in personnel management at the krasnoyarsk state agrarian university. *ANI: Economics and Management*, 1(26). 365-369. https://cyberleninka.ru/article/n/practical-experience-of-training-specialists-in-personnel-management-at-the-krasnoyarsk-state-agrarian-university

Gitman, M. B., Danilov, A. N., & Stolbov, V.Yu. (2014). Modern data analysis methods. *Open education*, 1(102), 24-31. https://openedu.rea.ru/jour/article/view/129

Gorlushkina, N.N., Kocyuba, I.Yu., & Khlopotov, M.V. (2015). Problems and methods of educational data mining to support decision making. *OTO*, 1, 472-482.

Haddouchane, Z. A., Bakkali, S., Ajana, S., & Gassemi, K. (2017). The application of the competency-based approach to assess the training and employment adequacy problem. *International Journal of Education (IJE).* 5(1), 1-18. https://doi.org/10.5121/ije.2017.5101

Kabakov, R.I. (2014) R in action. Analysis and visualization of data in R programm. *Moscow. DMK Press*, 588.

Kadralinova, M.T. (2015) Competence-based approach as solution of education quality. *Education through life: continuing education in the field of development study,* 13, 417-419.

Keim, D., Andrienko, G., Fekete, J-D., Gorg, C., Kohlhammer, J., & Melancon, G. (2008) Visual analytics: Definition, process, and challenges. *Human-Centered Issues and Perspectives, Springer,* 154-175, https://doi.org/10.1007/978-3-540-70956-5_7

Khan, M., & Khan, S. S. (2011) Data and information visualization methods, and interactive mechanisms: A survey. *International Journal of Computer Applications*, 34(1), 1-14.

Kulasegaram, K., & Rangachari, P. K. (2018) Beyond "formative": assessments to enrich student learning. *Advances in physiology education*, 42(1), 5-14, https://doi.org/10.1152/advan.00122.2017

Kumar, K., & Vivekanandan, V. (2018), Advancing learning through smart learning analytics: a review of case studies. *Asian Association of Open Universities Journal*, 13(1), 1-12, https://doi.org/10.1108/AAOUJ-12-2017-0039

Leontyev, V. V., Rebrina, F. G., Leontyeva, I. A., & Gafiyatullina E. A. (2016) Evaluation of the Development of Professional Competence in Undergraduates: Methodical Aspects. *International Journal of Environmental and Science Education*, 11(14), 6592-6602.

Linan, L. C., & Perez A. A. J. (2015) Educational Data Mining and Learning Analytics: differences, similarities, and time evolution. *International Journal of Educational Technology in Higher Education*, 12(3), 98-112, http://dx.doi.org/10.7238/rusc.v12i3.2515

Masino, S., & Nino-Zarazua, M. (2016) What works to improve the quality of student learning in developing countries? *International Journal of Educational Development*, 48, 53-65, https://doi.org/10.1016/j.ijedudev.2015.11.012

Munoz, D. R., & Araya, D. H. (2017) The challenges of competence-based assessment in the educational field. *Educacao e Pesquisa*, 43(4), 1073-1086, https://doi.org/10.1590/s1678-4634201706164230

Murumba, J., & Micheni, E. (2017) Big Data Analytics in Higher Education: A Review. *The International Journal of Engineering and Science (IJES),* 6(6), 14-21, http://dx.doi.org/10.9790/1813-0606021421

Oyerinde, Y. & Olu Bankole, F. (2015). Big Data in Education: Learning Analytics Techniques and Implementations for Predicting Students Academic Performance. https://doi.org/10.13140/RG.2.1.4004.3286

Occupational standards: 06 Communication, information and communications technology, Retrieved 08 April, 2020 from http://fgosvo.ru/docs/101/69/2/6

Rehman, S., & Mehmood, DR. A. (2014). Employability skills : the need of the graduates and the employer. *VSRD International Journals of Business and Management Research,* 4(IV), 133-138.

Ukuev, B.T. (2018) Experience of introduction of innovative methods of training of bachelors and masters in the field of information technologies. *Open Education*. 22(3), 83-90, https://doi.org/10.21686/1818-4243-2018-3-83-90

Utyomov, V.V., & Gorev, P.M. (2018) Development of educational systems on the basis of Big Data technology. *Concept*, 6, 449-461.

Vaitsis, C., Hervatis, V., & Zary, N. (2016) Introduction to Big Data in education and its contribution to the quality improvement processes. *Big Data on Real-World Applications*, 113, http://dx.doi.org/10.5772/63896

Zavadsky, J., Korenkova, V., Zavadska, Z., Kadarova, J., & Tucek, D. (2019) Competences in the Quality Management System Evaluation Based on the Most Worldwide Used Key Performance Indicators. *Quality-Access to Success*, 20(169), 29-41.