Marko MILOJKOVIĆ[1], Dragan STOJANOVIĆ[1], Natalija STOJANOVIĆ[1],
Aleksandra KŁOS-WITKOWSKA[2], Krzysztof WITOS[2], Marcin BERNAŚ[2],
Vasyl MARTSENYUK[2], Georgi DIMITROV[3], Dejan RANCIC[1],
Oleksiy BYCHKOV[4]

BADANIE POTRZEB I OCZEKIWAŃ W ZAKRESIE BIG DATA ABSOLWENTÓW INFORMATYKI MAGISTRÓW I ABSOLWENTÓW SYSTEMÓW I TECHNOLOGII INFORMATYCZNYCH

Streszczenie: Badania dotyczą absolwentów kierunków informatycznych, magistrów i absolwentów informatyki w zakresie systemów i technologii informatycznych. Ankieta została przeprowadzona on-line za pomocą narzędzi formularzy Google i została przeprowadzona w okresie od 1 września 2020 do 28 lutego 2021. Pytania zawarte w niniejszym opracowaniu miały na celu zbadanie potrzeb i oczekiwań związanych z Big Data, które mogą być wykorzystane jako wymagania przy opracowywaniu szkolenia Big Data.

Słowa kluczowe: Big Data, absolwenci IT, kompetencje, program nauczania

RESEARCH OF BIG DATA NEEDS AND EXPECTATIONS OF IT GRADUATES MASTERS AND IT ALUMNI IN INFORMATION SYSTEMS AND TECHNOLOGIES

Summary: The research is concerning IT graduates Masters and IT Alumni in Information Systems and Technologies. The survey was performed on-line using google forms tools and it was performed during a period from the 1'th of September 2020 to the 28'th of February 2021. The questions in this study aimed to research the needs and expectations related to Big Data that can be used as the requirements when developing Big Data training course

Keywords: Big Data, IT graduates, competence, curriculum

[1]University of Niš, Nis, Serbia: (marko.milojkovic, dragan.stojanovic, natalija.stojanovic, dejan.rancic)@elfak.ni.ac.rs
[2]Department of Computer Science and Automatics, University of Bielsko-Biala, Poland: (awitkowska, kwitos, mbernas, vmartsenyuk)@ath.bielsko.pl
[3]University of Library Studies and Information Technologies, Sofia, Bulgaria: g.dimitrov@unibit.bg
[4] Taras Shevchenko National University of Kyiv, Kiev, Ukraine: oleksiibychkov@knu.ua

1. Introduction

The designing innovative training course on Big Data is paying the attention in the series of works [1-4]. We also should take into account the experience of the existing offers regarding the large-scale data science [5-16]. This research was carried out in the context of project no. 2020-1-PL01-KA203-082197 entitled "Innovations for Big Data in a Real World". The survey was obtained by the IT graduates Masters and IT Alumni in Information Systems and Technologies. The survey was performed on-line using google forms tools and it was performed during a period from the 1'th of September 2020 to 28'th of February 2021. To obtain wide range of data, multiple choice question fields, with additional open question field were offered to overcome the shortcomings of narrowed answers suggestions. The survey contains both open and closed questions. To make a process of data collection un-biased no additional recommendation was added. No events were reported during that time that could influence the result. The survey data was presented in quantified form. The data was aggregated to the category based on syntactic analysis i.e., the difference in a form and not in a meaning. The open description was presented and analysed without modification.

1.1. Target

This study is a part of the research within IO1 in connection with the objectives of project 2020-1-PL01-KA203-082197 "Innovations for Big Data in a Real World" (BigWorld) under the Erasmus + program. The questions in this study aim to research the needs and expectations of IT graduates Masters and IT Alumni in Information Systems and Technologies.

1.2. Collection and analysis of data

The research was analyzed based on 631 questionnaires carried out in 4 countries participating in the project: Poland, Ukraine, Bulgaria Serbia.

3. Results

3.1. Nationality

The first group of three questions in the survey was of general purpose, with the aim to find out the basic information about questionees – their nationality, age and student status. The first question was about the nationality. Questionees were free to chose from partner countries (Poland, Bulgaria, Ukraine, Serbia) or to enter arbitrary answer.

Data description

Questionees were mostly residents of the participant countries of the project: Bulgaria 46.43%, Poland 25.04%, Ukraine 15.21% and Serbia 11.09%, as presented in the Figure 1 and the Table 1. There were also a few questionees from North Macedonia, Albania, Russia, Moldova, and Kazakhstan.
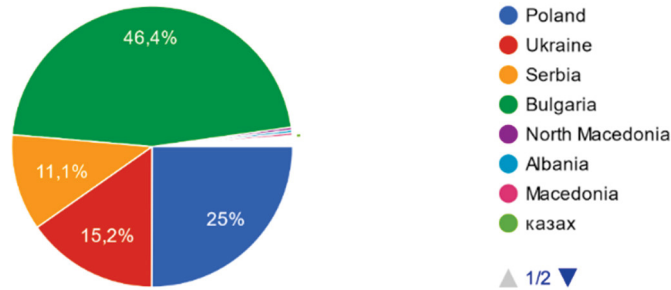
Figure 1. Nationality of respondents

Table 1. Nationality of respondents

Nationality	Total	Percentage
Bulgaria	293	46.43
Poland	158	25.04
Ukraine	96	15.21
Serbia	70	11.09
North Macedonia	8	1.27
Albania	2	0.32
Russia	2	0.32
Moldova	1	0.16
Kazakhstan	1	0.16

Discussion

Great response of IT graduates Masters and IT Alumni in Information Systems and Technologies (over 600 participants in the online survey), mostly from the partners' countries, demonstrates the interest of the target population in Big Data contents and a clear need for the Big Data courses.

3.2. Age

The second question was about the age of the questionees. Five ranges were offered: 17 to 20 year, 21 to 24 year, 25 to 34 year, 35 to 44 year, and 45 year and more.

Data description

Questionees were mostly 21 to 24 year old (37.56%), 17 to 20 year old (32.96%) and 25 to 34 year old (18.7%) as presented in the Fig. 2 and the Table 2. Smaller percentage of them were 35 to 44 year old (9.51%) and 45 year and more (1.27%).
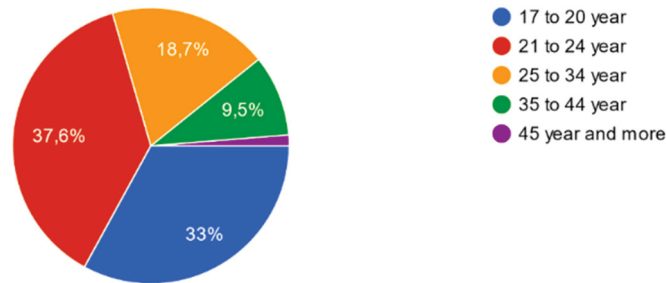
Figure 2. Age of respondents

Table 2. Age of respondents

Age	Total	Percentage
21 to 24 year	237	37.56
17 to 20 year	208	32.96
25 to 34 year	118	18.70
35 to 44 year	60	9.51
45 year and more	8	1.27

Discussion

Younger people are more interested in new technologies, especially those who are still at their universities (over 70% of respondents are younger than 24).

3.3. Students or graduates

The purpose of the third question was to establish whether questionees were students or graduates, so they had only these two options to chose from. Their choice directly affected the following questions. For instance, questions 04-13 were exclusive for students, while graduates were transferred directly to the fourth section (16th question).

Data description

Questionees were students in large majority - 552 (87.5%) while only 79 (12.5%) were graduates, as shown in the Figure 3.
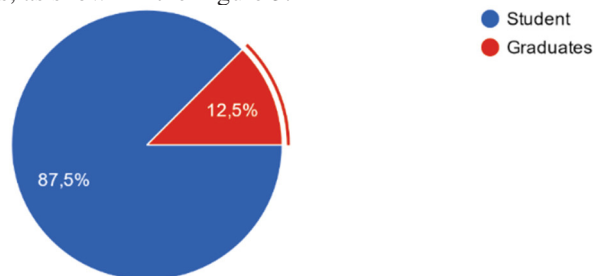

Figure 3. Distribution of students and graduates among the respondents

Discussion

As already stated, younger people are more interested in new technologies. Students clearly think that Big Data contents could be important for their further careers.

3.4. Education degree

Questionees that are students then had to answer to questions 4-13. Firstly, they were asked to state what education degree they are learning for. Offered answers were Specialist, Bachelor, Master, PhD, and Other.

Data description

Most of 552 students were from Bachelor Studies 337 (61.1%), Master Studies 131 (23.7%), Specialist Studies 59 (10.7%) and Engineer's Degree 16 (2.9%) as shown on the Figure 4.
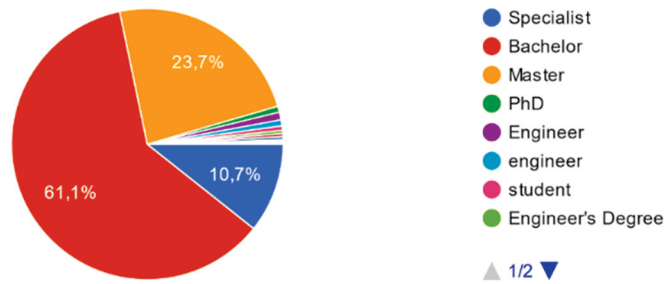

Figure 4. Education degrees of the respondents

Discussion

The interest of students in Bachelor and Master programmes for the field of Big Data in comparison with other types of studies at the academy, reveals us the target population among the IT community.

3.5. Specialty

The follow-up question was about the specialty of the students where they had to enter their answer in the free form.

Data description

The answers were very heterogeneous, but mostly some variations in the field of IT. Most common answers were: Information Technology with 28.26%, Informatics and Computer Science 18.66%, Software Engineering 12.86%, Computer Science 10.87%, Information Systems and Technologies 5.25%. The complete list is in Table 3 and Fig. 5.

Table 3. Specialties of the respondents

| Specialty	Total	Percentage
Information Technology	156	28.26
Informatics and Computer Science	103	18.66
Software Engineering	71	12.86
Computer Science	60	10.87
Information Systems and Technologies	29	5.25
Programming	18	3.26
Cybersecurity	15	2.72
Digital Marketing and Web Design	10	1.81
Data Analysis and Management	8	1.45
Information Security	8	1.45
Software Architectures and Quality Management	4	0.72
Software Development Techniques	4	0.72
Internet of Things	3	0.54
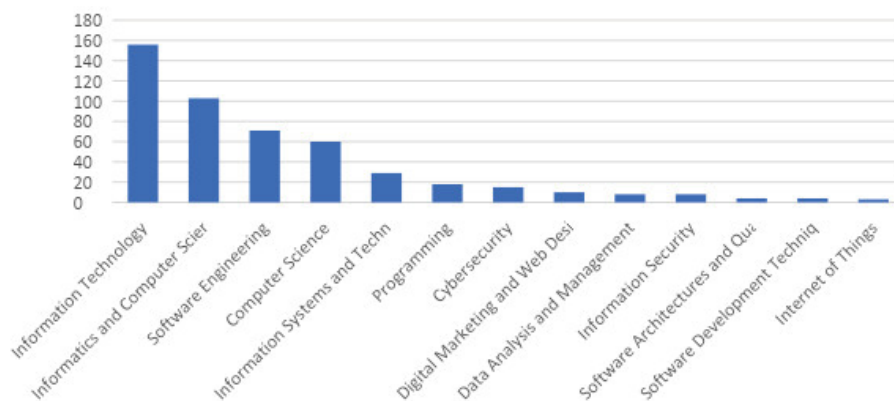
Figure 5. Specialties of the respondents

Discussion

Students of all kinds of IT studies showed interest in Big Data topics and courses, proving that the objectives of the project and target groups were set correctly.

3.6. Course

This question aims to find out the students' course number. Proposed answers were 1, 2, 3, 4, and the other (Fig. 6).

Data description

Most of the students chose option 1 (46.7%), 2 (21.4%), 3 (18.7%), and 4 (10%). Only 3.26% (18 students) entered some other answer (mostly 5 or 6).
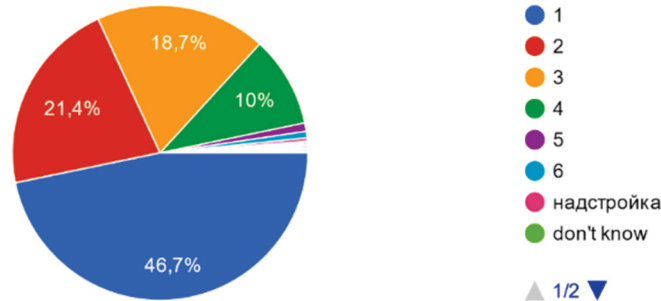
Figure 6. Year of the study of the respondent

3.7. Big Data courses

Students were asked whether they know any Big Data courses performed at their university. There were 3 proposed answers: Yes, I know a lot of them, Yes, I know only a few, and No (Fig. 7).

Data description

Students in large majority chose answers No (54.3%) and Yes, I know only a few (39.9%) while only 5.8% chose the answer Yes, I know a lot of them.

Do you know any Big Data courses performed at your university?
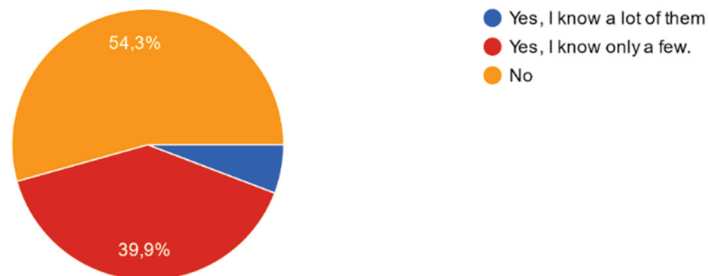552 одговора

Figure 7. Big Data courses at the university of the respondents

Discussion

The majority of students interested in the survey on the Big Data, but not knowing about any relevant course at their university, demonstrates the good foundation of iBigWorld project and the growing need for courses dealing with Big Data.

3.8. Preferred activities

The next question was what activities do students prefer to extend their knowledge in Big Data with offered answers: Lectures, Classes, Laboratories, Projects,

Apprenticeships, Student internships, Participation in the activity of students' scientific groups, Getting to know the results of research conducted at the University, Thematic courses / seminars / webinars, and Other. Students had an opportunity to choose multiple answers for this question (Fig. 8).
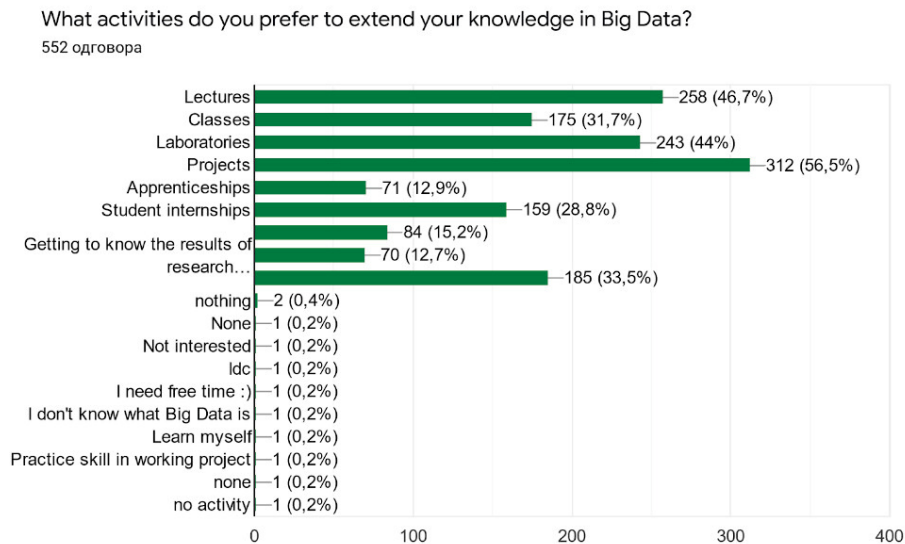
What activities do you prefer to extend your knowledge in Big Data?
552 одговора

Figure 8. Prefered activities of the respondents

Data description

The answers the students chose were the following: Projects 312 (56.5%), Lectures 258 (46.7%), Laboratories 243 (44%), Thematic courses / seminars / webinars 185 (33.5%), Classes 175 (31.7%), Student internships 159 (28.8%), Participation in the activity of students' scientific groups 84 (15.2%), Apprenticeships 71 (12.9%), Getting to know the results of research conducted at the University 70 (12.7%).

Discussion

Students showed interest in Big Data content and course and the developing course should lean heavily on practical implementation with laboratory work, students' projects and internships.

3.9. Future diploma thesis

The aim of this question was to find out how would the students like to use their Big Data research and skills for the aims of their future diploma thesis, with proposed answers: I would like to write a scientific paper, I would like to present my Big Data research and skills during seminar, I am not interested in Big Data activities, and other (Fig. 9).

Data description

The most of the students (239-43.3%) answered that they are not interested in Big Data activities, while 156 students (28.3%) would like to present their Big Data research and skills during seminar. The rest of the questionees (134-24.3%) mostly chose that they would like to write a scientific paper.

How would you like to use your Big Data research and skills for the aims of your future diploma thesis?

552 одговора

Figure 9. The usage of Big Data for diploma thesis

Discussion

Although the largest number of questionees answered that they are not interested in using Big Data in their future thesis, majority (more than 50%) responded positively, proving the necessity of introducing Big Data contents and courses through iBigWorld project.

3.10. Implementing Big Data subjects

The next question was what obstacles/issues do students see in implementing Big Data subject into a teaching program. Proposed answers were: Limitations resulting from the study program, Formal barriers in submitting a new form of classes, Inadequate or maladjusted laboratory/lecture room equipment, and Other, whereby students had the opportunity to choose multiple answers (Fig. 10).

Data description

266 students (52.5%) answered that they see Limitations resulting from the study program as an obstacle in implementing Big Data subject into a teaching programme, 198 (39.1%) gave the answer Formal barriers in submitting a new form of classes, while 111 (21.9%) stated Inadequate or maladjusted laboratory/lecture room equipment. Only 6% gave some other answer, mostly that they see no obstacles or that they don't know.
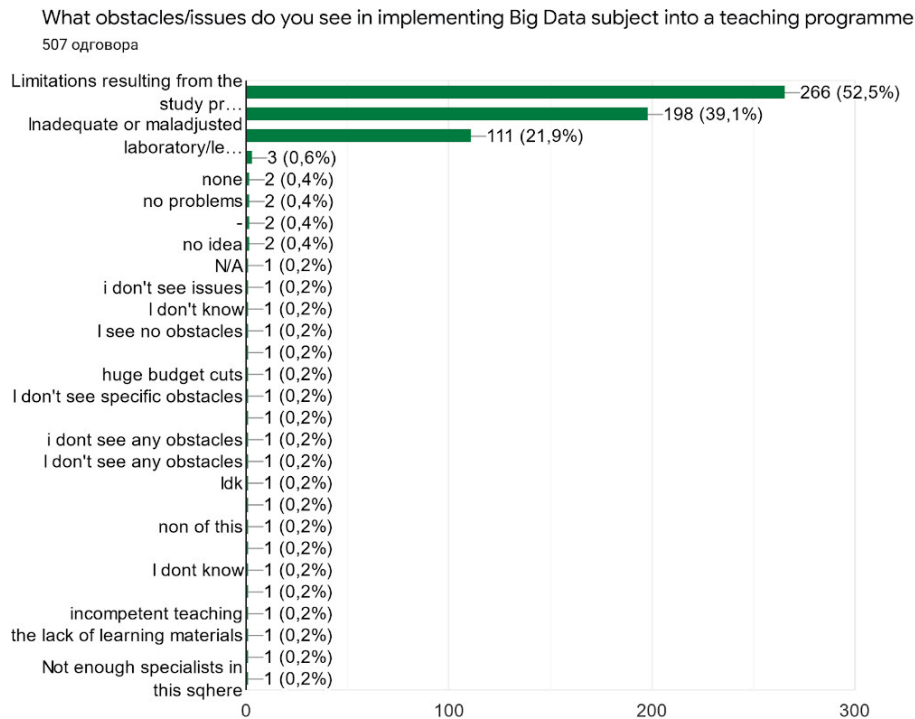
What obstacles/issues do you see in implementing Big Data subject into a teaching programme
507 одговора

Figure 10. Problems to solve when implementing Big Data course

Discussion

Partner countries of iBigWorld project should focus on implementing Big Data content into study programs by either introducing completely new courses or by innovating existing courses. Improving material component (equipment, laboratories) is secondary, but also important factor.

3.11. Building a Big Data website

The next topic the students were supposed to give their opinion was how do they rate the idea of building a website which will present results of Big Data research carried out by local University employees with 6 choices rating from Not at all important to Very important (Fig.11).

Data description

Majority of questionees (236 – 42.8%) had no specific opinion on this question, while similar number (42.2%) had opinion that building a website is important (Moderately important 29.9% and Very important 12.3%). Only 15% think that the website is not important (Slightly – 7.2%, Low – 5.3%, Not at all – 2.5%).

How do you rate the idea of building a website which will present results Big Data research carried out by local University employees?
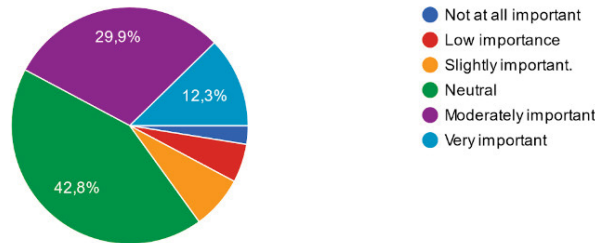552 одговора

Figure 11. Opinions on the building website

Discussion

Presenting of Big Data research in a form of website should be one of the most important goals for partner universities, because that kind of informing is the easiest way to disseminate the results and to motivate and include students and other target groups in Big Data topics.

3.12. Big Data issues

Another important question was whether the students are familiar with Big Data issues with four proposed answers: Yes, I know about it and I know how to use it, Yes, I have heard about the access but I have never used it, No, I have not heard about it and I would like to know how to use it, No, I have not heard about it and I am not interested in such access (Fig.12).
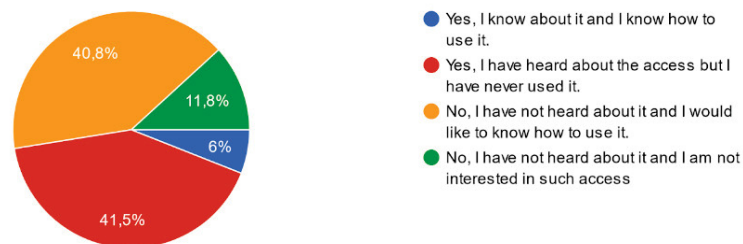
Are you familiar with Big Data issues?
552 одговора

Figure 12. Familiarities with Big Data

Data description

Similar number of students answered that they have heard about the access but never used it 229 (41.5%), and that they have not heard about it and would like to know how to use it 225 (40.8%). Much smaller percentage answered that they either not heard about it and would like to know how to use it (11.8%) or not heard about it and not interested in such access (6%).

Discussion

Large interest of the students in Big Data issues, despite their low level of insight adds to the project goal to identify the underrepresented skills, the rationale behind the phenomenon of talented people who lack the traditional credentials to land a good job and the Big Data with the most pressing needs.

3.13. Taking part in Big Data courses

The last and most important question in this group of questions was whether the students agree to take part in Big Data Courses. Proposed answers were: Disagree or Somewhat disagree (in these cases questionees were transferred to the fourth set of questions starting from 16th question), and Neutral, Somewhat agree, Agree (in these cases questionees were transferred to the third set of questions starting from 14th question) (Fig. 13).

Data description

Students were mostly Neutral regarding the option to take part in Big Data courses (230 – 41.7%), Somewhat positive (155 – 28.1%) or Positive (109 – 19.7%). Minority of the students were Negative or Somewhat negative (both 29 – 5.3%).
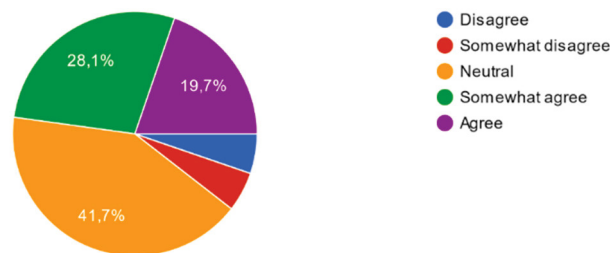

Figure 13. Engagement in Big Data course

Discussion

Having almost all of the IT students interested or neutral regarding studying Big Data course, is in line with assumption of the project that there is a gap in southern Europe regarding the Big Data and other digital skills and a need for new initiatives to ensure that these skills are adequately promoted in the curriculum, in teacher development, in assessment practices and in learning content.

3.14. Big Data content

Students which stated that they would like to take part in Big Data Courses in question 13, had the opportunity to answer the third group of questions Participation in Big Data training consisting of only two questions (14 and 15). First of those two questions was what would they like to do, in the context of Big Data, within their classes as a part of Big Data programme. Offered answers were: I am interested in collecting

data from different sources (web, social networks etc.), I am interested in doing Big Data analytics/ machine learning, I am interested in presenting data processed (uploading results into websites and more), and Other (Fig. 14).

Data description

Almost the same number of students are interested in tow aspects of Big Data science: analytics/ machine learning (202 – 40.9%) and collecting data from different sources (195 – 39.5%). Much smaller percentage (17.2%) chose that they are interested in presenting data processed.

What would you like to do, in the context of Big Data, within your classes as a part of Big Data programme:
494 одговора

Figure 14. Opinions on the content of the course

Discussion

Obtained results can be used for designing the attractive Big Data course heavily based on techniques of collecting data from different sources as well as on analytics of those data, i.e. machine learning.

3.15. Previous Big Data courses

The other question related to the Participation in Big Data training was whether students previously participated in classes which were based on Big Data (Fig.15).

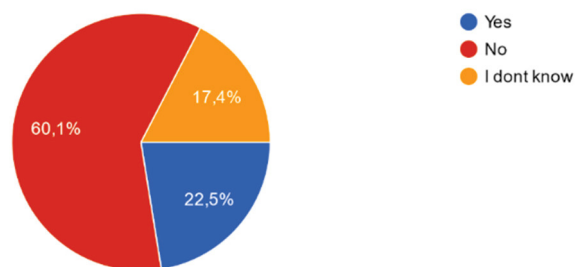Have you participated in classes which were based on Big Data
494 одговора

Figure 15. Previous Big Data courses

Data description

Answers were straightforward: 297 students (60.1%) haven't participated in classes related to Big Data, 111 (22.5%) did, and 86 (17.4%) and not sure.

Discussion

Only one fifth of the students which already participated in classes based on Big Data (and we must have in mind that this question was answered exclusively by the students interested in Big Data content) demonstrates the lack and the need for offering adequate Big Data courses at partner countries' universities.

3.16. Current job

The question 16 in the online survey was important for the future course of the survey. Questionees were asked whether they currently have a paid job. If their answer was Yes they had to answer questions from sections 5 (Work experience, starting from question 17) and 6 (Workers in IT industry, starting from question 20), while those answering No skipped all consequent questions until section 7 (Important competencies, starting from question 24).

Data description

As can be seen from the following Figure, questionees were almost perfectly divided: 321 (50.9%) answered that they currently have paid job, while 310 (49.1%) answered negatively (Fig. 16).
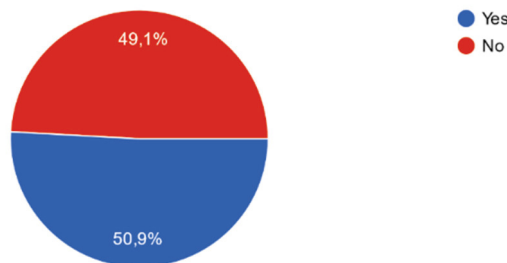
Are you working now?
631 одговор

Figure 16. Current employment

Discussion

This question was important only for the selecting further questions in the survey, but it is interesting that almost the same number of questionees are currently working and are without a paid job. This proportion allows needed diversity for the survey analysis.

3.17. Work sector

Section 5, related to the work experience, consists of three questions (17-19). First one was about the sector in which the questionee is currently working with offered answers: Private sector, Public sector, Non-profit organizations, and Other.

Data description

Large majority of questionees stated that they are currently working in Private sector (231 – 66.4%), while 87 (27.1%) are working in Public sector. Only 11 questionees (3.4%) work in Non-profit organizations. 10 questionees provided some other answer (Fig. 174).

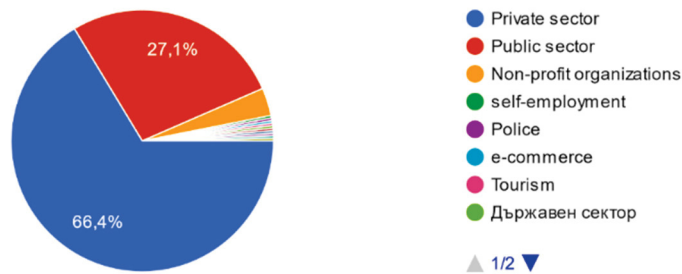In which sector do you work?

321 одговор

Figure 17. Branch of the employment

Discussion

Designed Big Data courses should be tailored particularly toward people working in the private sector (by content, time of lectures...). These people are mostly motivated to improve their competencies and to learn more about Big Data.

3.18. Work experience

The aim of the next question was to find out for how long employed questionees work in their present work places with four ranges proposed: less than 1 year, between 1 and 3 years, between 3 and 5 years, and over 5 years.

Data description

As can be seen from the Figure below, questionees are mostly in the beginning of their working careers with most of them working in the current position less than 1 year (37.4%) or between 1 and 3 years (34%). 11.5% are working between 3 and 5 years while 17.1% work over 5 years in their current position (Fig. 18).

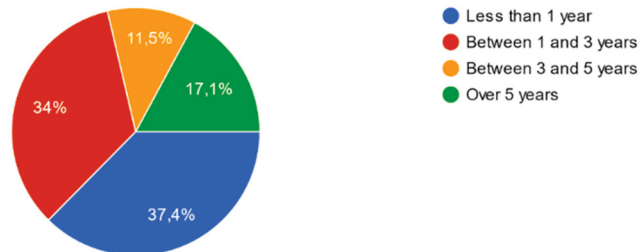How long have you been working in this position?
321 одговор

Figure 18. Work experience

Discussion

Proposed Big Data courses should target younger people with less working experience because they are more motivated and still interested in learning new topics, for which they feel that could help them improve their careers.

3.19. Work field

The final question in this section aims to reveal the work field of the questionees. Offered answers are: IT industry, Manufacture, Economics (accounting, own business), Administration, Marketing, Banking, Security and defence activities, and Other. Those who choose IT industry, Banking or Security and defence activities continue with the questions from the next section (section 6 with questions 20-23) while the rest proceed with section 7 (beginning with question 24).

Data description

As expected, large majority of questionees work in IT industry – almost 60%. Large groups are also Administration (8.1%), Marketing (7.17%), Security and defence activities (4.36%), Economics (4.05%), Manufacture (2.49%), Catering (2.18%), Banking (1.87%), and Education (1.56%), as shown in Figure 19 and Table 4.

Table 4. Work fields of the respondents

Work field	Total	Percentage
IT industry	191	59.50
Administration	26	8.10
Marketing	23	7.17
Security and defense activities	14	4.36
Economics (accounting, own business)	13	4.05
Manufacture	8	2.49
Catering	7	2.18
Banking	6	1.87
Education	5	1.56

Tourism	2	0.62
Logistics	2	0.62
Aviation	2	0.62
Sport	1	0.31
Healthcare	1	0.31
Media	1	0.31
Judiciary	1	0.31
Pharmacy	1	0.31
Construction	1	0.31
Agriculture	1	0.31

What field do you work in?

321 одговор

Figure 19. Work fields

Discussion

Proposed courses in Big Data should be designed toward more professional audience because it is much lesser chance that someone other than IT specialists from would be interested in Big Data content.

3.20. Job inside the IT industry

Section 6 of the survey was reserved only for Workers in the IT industry, and it consists of four questions (20-23). First question was to state concrete job inside IT industry, like: Network construction and service, Software developer, Graphic design, Telecommunications, Data Analytics, Data Science, Support, Administrator, Big Data Developers, Database Administration, and Other.

Data description

It can be seen from the Figure 20 and Table 5 that jobs inside IT industry are very diverse. Most of IT workers work as Software developers (51.43%), Support (11%), Administrators (8.1%) or in Network construction and service (4.76%).

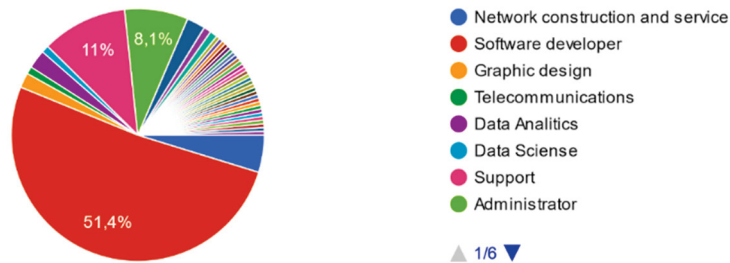What is your job in the IT industry?
210 одговора

Figure 20. Jobs in IT

Table 5. IT positions

Job position	Total	Percentage
Software developer	108	51.43
Support	24	11.43
Administrator	17	8.10
Network construction and service	10	4.76
Quality Assurance	7	3.33
Data Analytics	6	2.86
Database Administration	5	2.38
Graphic design	4	1.90
Business analyst	3	1.43
Telecommunications	2	0.95
Data Science	2	0.95
SAP Consultant	2	0.95

Discussion

Designed Big Data courses must be heavily oriented toward Software developers because they are making the majority of the target group for developing digital competencies and skills in the area of Big Data.

3.21. Required education

The follow-up question for workers in IT sector was what kind of education is required for an appointment to their position: Bachelor, Vocational high school, Professional specialized course, No requirements, or Other.

Data description

Out of 210 workers in IT industry, 126 (60%) stated that for their position Bachelor degree is required. 41 (20%) works at positions where there is No requirements, while 28 (13.3%) needed to finish Professional specialized course. 11 employees (5.2%)

work with Vocational high school. Two questionees answered that they needed Master degree for their positions, while one answered Bachelor or proven professional record (Fig. 21).

What education is required for an appointment to your position?
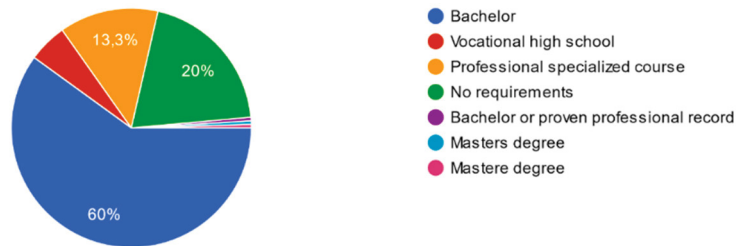210 одговора

Figure 21. Education needed

Discussion

During designing Big Data trainings, project partners must have in mind that workers in IT sector either work at positions where there are no requirements regarding their education or bachelor diploma is needed (together 80%). These data are consistent with the lack of IT experts in the market. There are almost no workers who must have master degree. So, the most logical choice to place Big data courses would be at Bachelor studies or as part some specialization programmes.

3.22. Experience in Big Data

Next question aims to find out whether IT workers have experience in Big Data field and for how long. Offered answers were: No experience, Short practice up to one year, From 1 to 3 years, From 4 to 5 years, over 5 years.

Data description

Surveyed IT workers mostly (71%) have no experience in Big Data field or have Short practice up to one year (22.9%). Only 4.8% have experience From 1 to 3 years, 1% from 4 to 5 years and 0.5% over 5 years (Fig. 22).
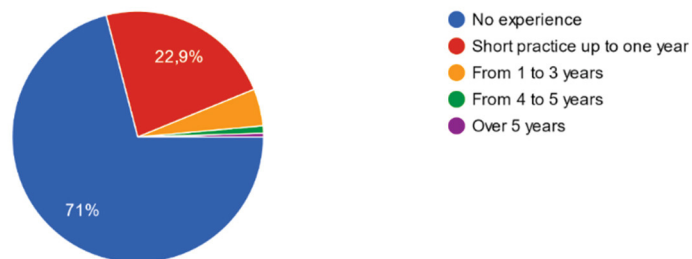
What is Your experience in BigData field?
210 одговора

Figure 22. Experience in Big Data of the respondence

Discussion

Large majority of workers in IT sector either have no experience or very limited experience (up to one year) in Big Data field meaning that designed courses have to be at the basic level providing all the necessary knowledge for dealing with Big Data.

3.23. Use of Big Data

Last question in section dedicated for IT workers was whether they use Big Data at their work with four options: Yes, to a small extent, Yes, to a large extent, No, I did not search such tools.

Data description

Vast majority of 210 surveyed IT workers do not use Big Data at their work (142 or 67.6%). 49 of them (23.3%) use it, but only to a small extent. Only 10 (4.8%) use Big Data to a large extent (Fig. 23).
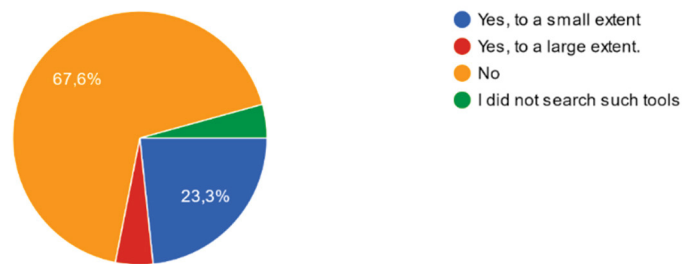
Do you use BigData at your work?
210 одговора

Figure 23. Usage of Big Data by respondents

Discussion

Two thirds of the respondents do not use Big Data at their work meaning that they either do not need it or they are not aware that these tools could ease their everyday work. Both to them and to those IT workers which use Big Data tools to some extent, Big Data contents would be valuable for their future career development.

3.24. Importance of Big Data issues

Question 24 starts a next set of questions (section 7) consisting of 6 question related to Important competencies for Big Data. This question is also a hub where all the questionees gathers again (all 631 of them). The first question of the group was: Do you think Big Data issues are important for your future career? Offered answers were: Not at all important, Low importance, Slightly important, Neutral, Moderately important, and Very important.

Data description

37.9% of questionees took Neutral stance toward Big Data competencies. 26.8% think that Big Data competencies are Moderately important for their future career, 11.3% consider them Slightly important, while 11.1% think that Big Data competencies are Very important. Only 7.9% consider them of Low importance and 5.1% Not important at all (Fig. 24).

Do you think Big Data issues are important for your future career?
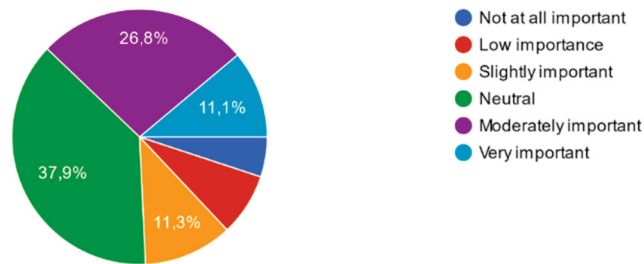
631 одговор

Figure 24. Opinions on the importance of Big Data

Discussion

Only 13% of respondents answering that Big Data issues are not important to their job prove the foundation of iBigWorld project and importance of introducing structured trainings in Big data topics.

3.25. Important competencies

This question was different from the rest because questionees had to select 10 competencies from the broader list (31 competencies in total), which they consider most important for the employment in the organization in which they work (Table 6).

Table 6. Study of the significance of the competencies

Competency	Total	Percentage
Ability to work in a team	431	68.30
Ability to plan and manage time	427	67.67
Ability to communicate in a second (foreign) language	415	65.77
Capacity to learn and stay up-to-date with learning	353	55.94
Ability to identify, propose and resolve problems	329	52.14
Ability to apply knowledge in practical situations	306	48.49
Capacity to generate new ideas (creativity)	304	48.18
Ability to be critical and self-critical	267	42.31
Ability to search, process and analyze information from a variety of sources	262	41.52
Ability to work autonomously	243	38.51

Ability to make reasoned decisions	234	37.08
Ability to use information and communication technologies	232	36.77
Commitment to tasks and responsibilities	230	36.45
Ability to design and manage projects	222	35.18
Ability to adapt to and act in new situations and cope under pressure	204	32.33
Ability to communicate both orally and through the written word in first language	193	30.59
Knowledge and understanding of the subject area and understanding of the profession	192	30.43
Ability to communicate with people who are not experts in the subject area	183	29.00
Ability to work in an international context	173	27.42
Ability for abstract and analytical thinking, and synthesis of ideas	168	26.62
Ability to interact with others in a constructive manner, even when dealing with difficult issues	166	26.31
Ability to motivate people and move toward common goals	163	25.83
Ability to evaluate and maintain the quality of work produced	159	25.20
Commitment to health, well-being and safety	114	18.07
Ability to act on the basis of ethical reasoning	110	17.43
Ability to undertake research at an appropriate level	105	16.64
Ability to interact constructively with others regardless of background and culture and respecting diversity	90	14.26
Commitment to conservation of the environment	81	12.84
Ability to act with social responsibility and civic awareness	80	12.68
Ability to take the initiative and to foster the spirit of entrepreneurship and intellectual curiosity	75	11.89
Ability to show awareness of equal opportunities and gender issues	63	9.98

Data description

Results of the survey are given in the table below. Few competences stand out in the sense that large majority of the workers find them important: Ability to work in a team (68.3%), Ability to plan and manage time (67.67%), Ability to communicate in a second (foreign) language (65.77%). The following competences were also popular: Capacity to learn and stay up-to-date with learning (55.94%), Ability to identify, propose and resolve problems (52.14%), Ability to apply knowledge in practical situations (48.49%), Capacity to generate new ideas (creativity) (48.18%). Other competences were chosen less frequently and the following competences were the least important for our questionees: Commitment to conservation of the environment (12.84%), Ability to act with social responsibility and civic awareness (12.68%), Ability to take the initiative and to foster the spirit of entrepreneurship and intellectual curiosity (11.89%), Ability to show awareness of equal opportunities and gender issues (9.98%).

Discussion

Developing of digital and soft skill becomes more and more important in today's job market. Young people recognize their importance as well as their employers. Any designed training should also focus on developing soft skills, especially those related to teamwork, communication and time management.

3.26. Competencies of specialist in Big Data

Question 26 was first in a several questions where the task was to rate certain competencies with 1 to 5 rates. This question has the aim to graduate competencies (in terms of their importance) which the specialist in Big Data should have. Offered competencies were the following four:
1. ability to process large volumes of data using hierarchical storage, hashing and filtering,
2. ability to select the efficient algorithm to Big Data, which takes under consideration its scale,
3. ability to select appropriate sampling and filtering method for given Big Data analysed case,
4. ability to tackle with concurrency / parallelism problems of Big Data scale.

Data description

As can be seen from Figure 26 below, questionees rates all four competencies related to Big Data relatively high, giving them mostly high rates 3-5.

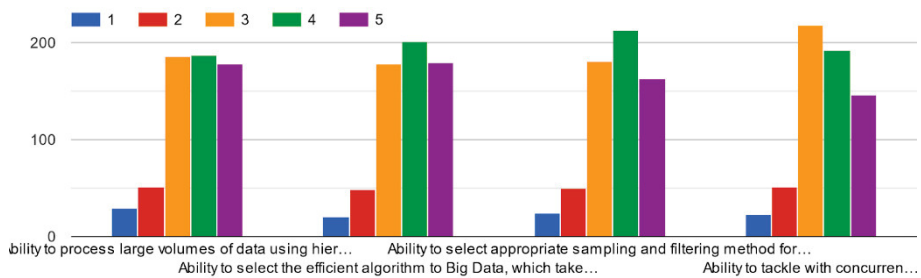What competencies (in terms of their importance) should have a specialist in Big Data? Put a point from 1 to 5

Figure 26. Professional competencies

Discussion

Any designed course in Big Data has to take into account the main topics related to this area: processing large volumes of data using hierarchical storage, hashing and filtering; selecting the efficient algorithm to Big Data, which takes under consideration its scale; selecting appropriate sampling and filtering method for given

Big Data analysed case; tackling with concurrency / parallelism problems of Big Data scale.

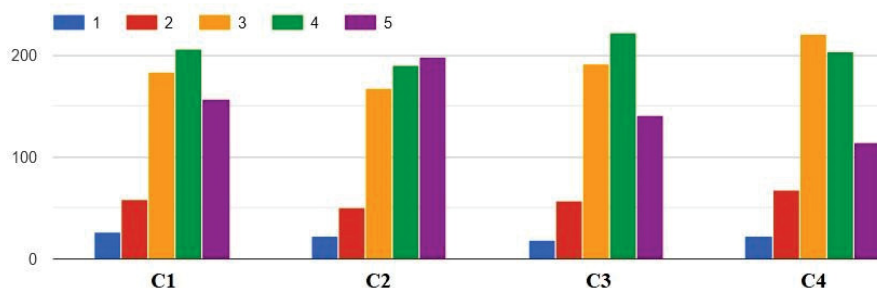### 3.27. Related competences of specialist in Big Data

The follow-up question was What related competences (in terms of their importance) should have a specialist in Big Data? Questionees also had to put their rates from 1 to 5 next to each of 12 related competencies:

1. Obtain information from existing sources (streaming data/ historical ones/ applications logs/ open-source databases).
2. Effectively use variety of data analytics techniques (Machine Learning, Data Mining, Prescriptive and Predictive Analytics).
3. Apply quantitative techniques (statistics, time series analysis, optimization, and prediction).
4. Process heterogeneous data (natural language, visual objects, data, text and other).
5. Visualise results of analysis of Big Data.
6. Deploy solution (merging data collection, storage, analysis and visualisation).
7. Using wide range of Big Data analytics platforms.
8. Develop and operate large scale data storage (e.g., Data Lakes, Hadoop and others).
9. Apply data security mechanisms and controls at each stage of the data processing.
10. Design, build, operate relational and nonrelational databases (SQL and NoSQL).
11. Process large dataset (i.e., ETL, OLTP, OLAP).
12. Ensure data quality, accessibility, interoperability, compliance to standards, and publication (data curation).

Data description

Ratings distribution for all 12 competencies can be seen in Figures 27 below. Competencies are labelled as C1-C12. We can see similar trends as in the previous question where questionees rates all the competencies relatively high (mostly with rates 3-5).

What related competences (in terms of their importance) should have a specialist in Big Data? Put a point from 1 to 5

What related competences (in terms of their importance) should have a specialist in Big Data? Put a point from 1 to 5
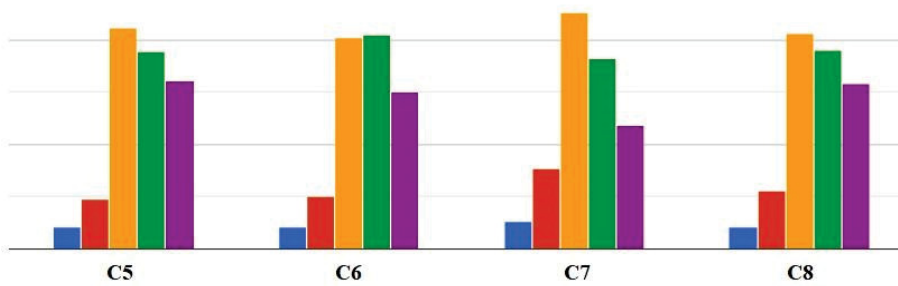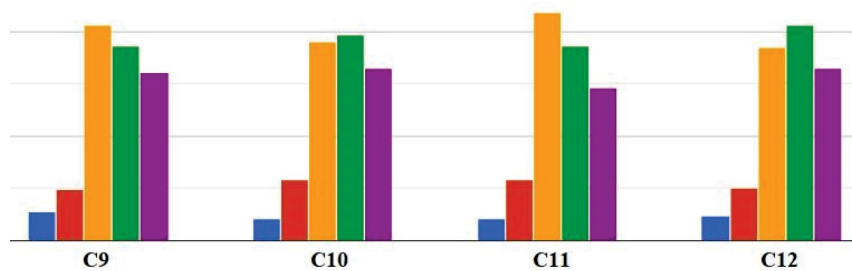

What related competences (in terms of their importance) should have a specialist in Big Data? Put a point from 1 to 5

Figure 27. Related competencies

Discussion

Most valued competences which a specialist in Big Data should have and which should be taught during Big Data trainings is Effectively use variety of data analytics techniques (Machine Learning, Data Mining, Prescriptive and Predictive Analytics). Analysis of other survey questions also demonstrate that Machine Learning and Data Mining are the most popular aspects of Big Data and data analytics.

3.28. Big Data tools

Aim of this question was to find out how the questionees rate the need of possessing knowledge on certain Big Data tools. Offered ratings ranged from 1 (low importance) to 5 (high importance).
   1. Big Data and distributed computing tools (Spark, MapReduce, Hadoop, Mahout, Lucene, NLTK, Pregel, etc.).
   2. Big Data Analytics platforms (Hadoop, Spark, Data Lakes, others).
   3. Real time and streaming analytics systems (Flume, Kafka, Storm).
   4. Hadoop Ecosystem/platform.
   5. Azure Data Analytics platforms (HDInsight, APS and PDW, etc).

6. Amazon Data Analytics platform (Kinesis, EMR, etc).
7. Other cloud-based Data Analytics platforms (HortonWorks, Vertica LexisNexis HPCC System, etc).
8. Cognitive platforms (such as IBM Watson, Microsoft Cortana, others).
9. Kaggle competition, resources and community platform.

Data description

Ratings distribution for all 9 Big Data tools can be seen in Figures 28. Tools are labelled as T1-T9.
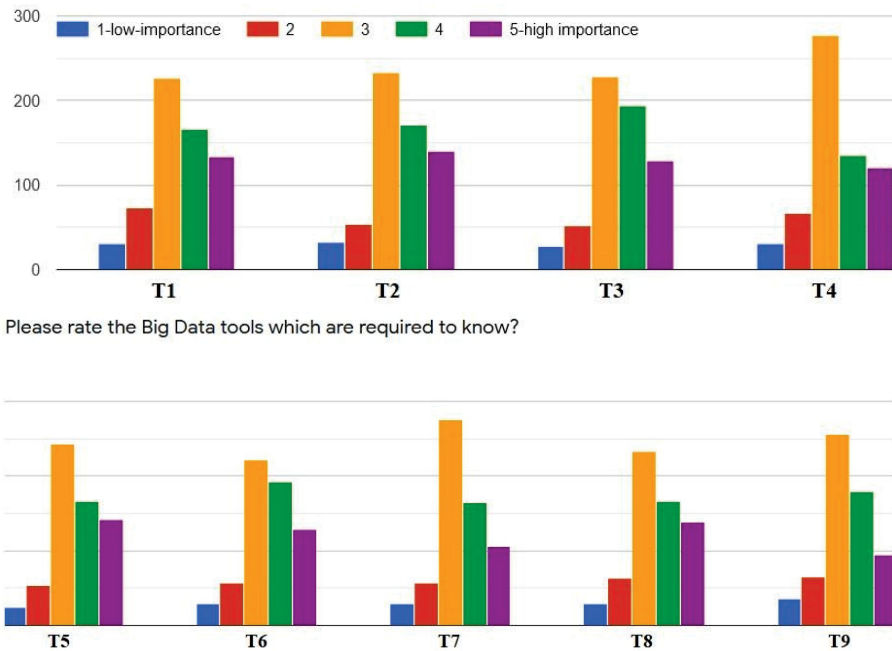
Figure 28. Importance of Big Data tools

Discussion

Answers to this question can help us design popular Big Data courses in regard of practical Big Data tools to be used to assist theoretical lectures, among them Big Data Analytics platforms and Real time and streaming analytics systems.

3.29. Academic/analytical Big Data Competencies

The last question in this group (Important competencies for Big Data) related to the rating the competencies which academic/analytical Big Data employee should have

(offered ratings were the same, ranging from 1 (low importance) to 5 (high importance)):

1. Ability to perform simulations and experiments.
2. Ability to verify results with statistical tools.
3. Ability to carry out feasibility studies on new technologies, methods, and standards that could be of use to the organization.
4. Ability to innovate and modify methods and approaches used in the organization.
5. Ability to write research or technical papers on the results of work.
6. Ability to apply modern methods of psychology and pedagogy in everyday work.
7. Ability to patent inventions and technical innovations, to perform standardization of developed systems and processes.
8. Ability to ensure and manage copyright protection of software products and to carry out their price evaluation.

Data description

Ratings distribution for all 8 competencies can be seen in Figures 29. Competencies are labelled as C1-C8.
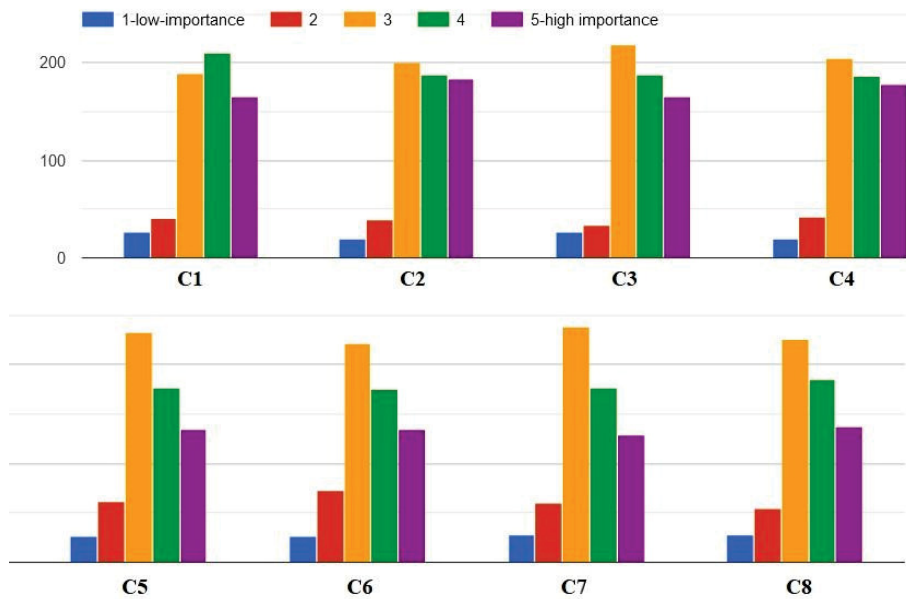

Figure 29. Academic competencies

Discussion

This question gives us the clear direction which future Big Data trainings should take. Workers in IT sector find the following competencies most important for their future

careers: verifying results with statistical tools; innovating and modifying methods and approaches used in the organization; and ability to perform simulations and experiments.
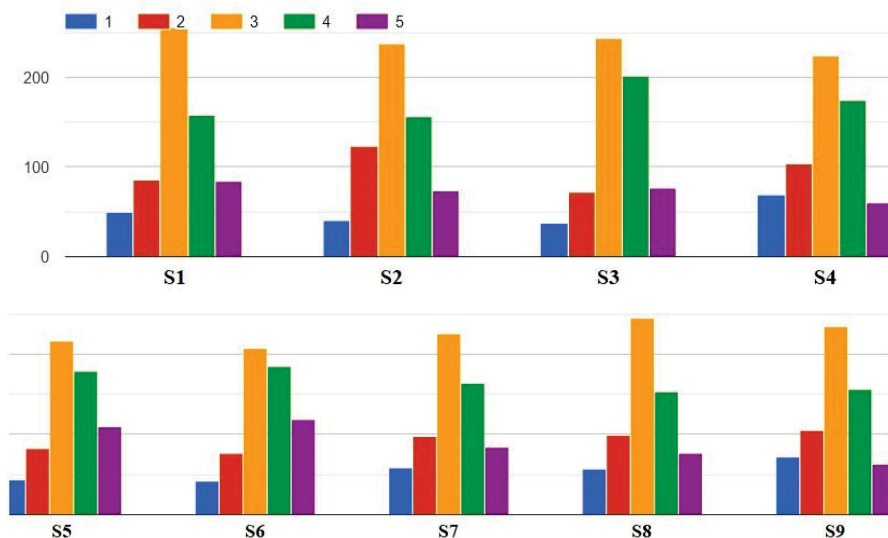
3.30. Competency training

The last set (section) of questions, labelled Your completed competency training, consists of five questions. First one was about evaluating competency training where the respondents had to rate from 1 to 5 the following statements:
1. Mastering the techniques of scientific work.
2. Mastering the ability to write scientific texts.
3. Individual professional and educational consulting.
4. Mastering Business Analysis methods.
5. Acquiring skills in managing IT projects.
6. Professional knowledge and skills in software development.
7. Professional knowledge and skills in the development of IS.
8. Acquiring the skills of system administration.
9. Mastering the ability to use industrial architecture in the development of IS.
10. Mastering software testing technology and IS.
11. Acquiring skills systems analyst.
12. Development of logical, algorithmic, systems thinking.
13. Ability to adapt to changes in the IT market.

Data description

Ratings distribution for all 13 statements can be seen in Figures 30. They are labelled as S1-S13.

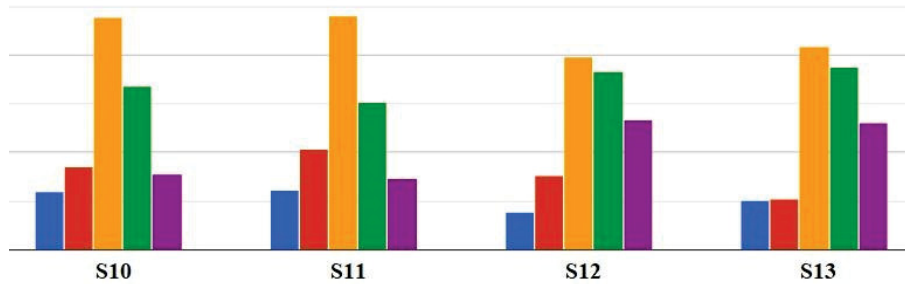Please evaluate your completed competency training. Put a point from 1 to 5

Figure 30. Competency training

Discussion

Evaluating competency training demonstrated that IT workers highly value both soft competencies like Development of logical, algorithmic, systems thinking and Ability to adapt to changes in the IT market, as well as professional skills like Knowledge and skills in software development and Acquiring skills in managing IT projects.

3.31. Encountered difficulties

This question aims to discover: What difficulties have questionees encountered in their employment? They could choose multiple answers from offered: At my specialty sites offer little, Knowledge is required, which I do not have, Offered places does not correspond with the profile of my training, Another level of higher education required, Need graduates with other specialties, Need competence, which I do not have.

Data description

As shown in the Figure 31, most of the respondents answered that they encountered situations in their employment when Knowledge was required, which they did not posses (260 – 41.2%) or Competence was needed, which they did not have (217 – 34.4%). Smaller number encountered Requirements of another level of higher education (151 – 23.9%), Offered places did not correspond with the profile of their training (144 – 22.8%), and At their specialty sites offer little (141 - 22.3%). Only 105 (16.6%) encountered situation where Graduates with other specialties were needed.

Discussion

Difficulties that IT workers encountered during their career once again confirm the proclaimed goals of iBigWorld project. Namely, IT workers answered in large majority that they encountered situations where Knowledge was required, which they did not posses or Competence was needed, which they did not have. Both problems should be tackled by designing adequate Big Data trainings in cooperation with business sector.
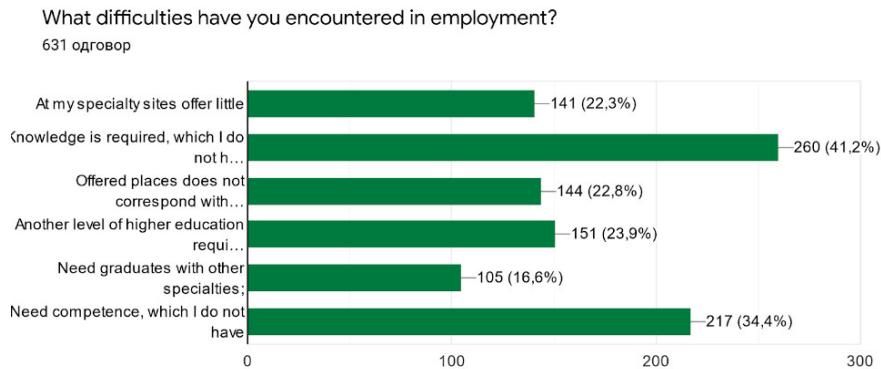
What difficulties have you encountered in employment?
631 одговор

Figure 31. Possible difficulties

3.32. Professional prospects

Respondents were then questioned to evaluate their professional prospects in the future. They were offered to put a point from 1 to 5 to two important questions: Reliability employment and Possibility of professional conversion.

Data description

The obtained ratings for two questions are given in Figure 32.

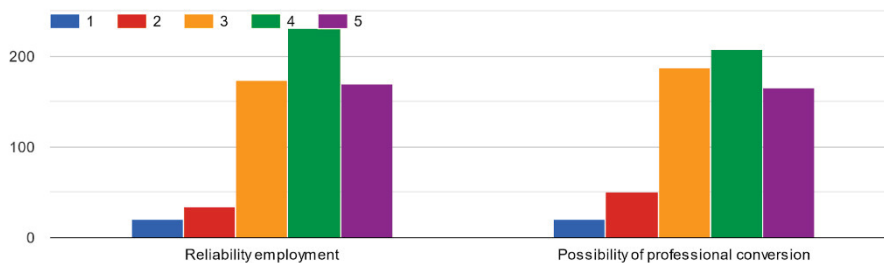Evaluate your professional prospects in the future? Put a point from 1 to 5

Figure 32. Professional prospects

3.33. Preferred profession

The next question was to Select a profession for which the respondents are working or want to work (in accordance with the list of the European framework of IT competences). They could select multiple answers from the following: Service Desk Agent, Technical Specialist, ICT Trainer, Business Analyst, Business Information Manager, Project Manager, ICT Consultant, Digital Media Specialist, Test Specialist, Software Developer, Systems Architect, Systems Administrator, Information System Developer, ICT Operations Manager, ICT Security Specialist, Enterprise Architect, Quality Assurance Manager, Chief Information Officer (CIO), Network Specialist,

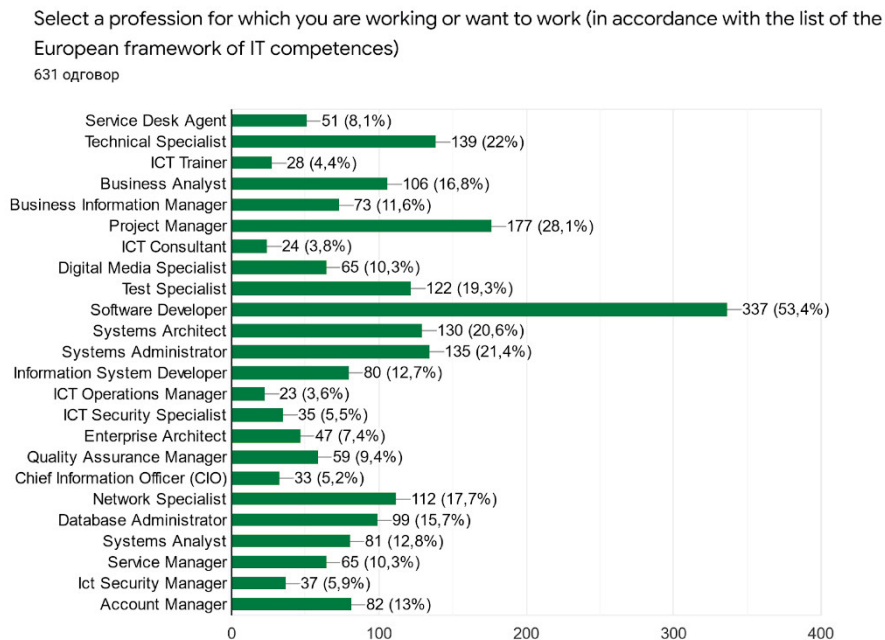Database Administrator, Systems Analyst, Service Manager, ICT Security Manager, Account Manager (Fig. 33).

Select a profession for which you are working or want to work (in accordance with the list of the European framework of IT competences)

631 одговор

Figure 33. Preferrable positions

Data description

Overwhelming majority selected option Software Development (337 – 53.4%). Popular options are also: Project Manager (177 – 28.1%), Technical Specialist (139 – 22%), Systems Administrator (135 – 21.4%), and Systems Architect (130 – 20.6%). Least popular professions are ICT Trainer (28 - 4.4%), ICT Consultant (24 – 3.8%), and ICT Operations Manager (23 - 3.6%).

Discussion

Big Data trainings and courses and guidelines for their design have to take into account desirable professions for which the respondents are working or want to work (in accordance with the list of the European framework of IT competences) and to be tailored specifically around them: Software Development, Project Manager, Technical Specialist, Systems Administrator, and Systems Architect.

3.34. Work environment

The last question was to describe the place of labour activity and work environment. Questionees could select multiple answers from the following:
  I work above projects
  Specialists of different profiles are in collective, which I work in
  I regularly co-operate with colleagues which have other profile of preparation

My work is often estimated
Success is acknowledged
I count on own forces at the decision of problem
An innovative climate dominates
I must think wider than scopes of my specialty often
My initiative is valued
My tasks change often
I have possibility to make decision in my sphere of activity
My work requires loneliness
My work is good time certain
I organize my work Independent
My work is related to international cooperation
My suggestions on perfection in earnest are checked up
I need a foreign language in everyday work
A large value gets to further preparation and in-plant training
A co-operative atmosphere dominates

Data description

Answers were very various, but several options stand out: My tasks change often (209 – 33.1%), Success is acknowledged (202 – 32%), I need a foreign language in everyday work (186 – 29.5%). Statements which least describe the place of labour activity and work environment are: My suggestions on perfection in earnest are checked up (46 – 7.3%) and A large value gets to further preparation and in-plant training (44 – 7%) (Fig. 3).

How did you describe the place of labor activity and your work environment?
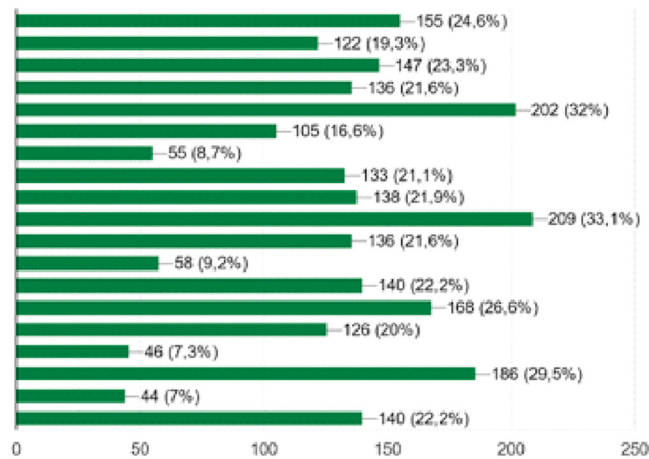631 одговор

Figure 34. Work environment

Discussion

Work environment of IT workers can be important for designing Big Data courses and pushing adequate competencies. Survey showed that workers tasks change often, and they need to know foreign language in everyday work but, in turn, their success is acknowledged.

4. Conclusions

So, the research was concerning IT graduates Masters and IT Alumni in Information Systems and Technologies. The questions in this study enabled us to research the needs and expectations related to Big Data that can be used as the requirements when developing Big Data training course.

REFERENCES

1. TINOCO-MARTÍNEZ S.R., FERREIRA-MEDINA H., CENDEJAS-VALDEZ J.L., HERNÁNDEZ-RENDÓN F., FLORES-MONROY M.M., GINORI-RODRÍGUEZ B.H.: How to Improve the Teaching of Computational Machine Learning Applied to Large-Scale Data Science: The Case of Public Universities in Mexico (2023) Lecture Notes in Networks and Systems, 544 LNNS, pp. 1-15. DOI: 10.1007/978-3-031-16075-2_1.
2. HAENLEIN M., KAPLAN A.: A brief history of artificial intelligence: On the past, present, and future of artificial intelligence (2019) California Management Review, 61 (4), pp. 5-14. DOI: 10.1177/0008125619864925.
3. DENG L., YU D.: Deep learning: Methods and applications (2013) Foundations and Trends in Signal Processing, 7 (3-4), pp. 197-387. DOI: 10.1561/2000000039.
4. HUANG L., MA K.-S.: Introducing machine learning to firstyear undergraduate engineering students through an authentic and active learning labware (2018) 2018 IEEE Frontiers in Education Conference (FIE), pp. 1-4.
5. Webpage: *https://imp.i384100.net/zavqGm* (Last accessed 16-10-2022).
6. Webpage: *http://datacamp.pxf.io/e4NgD6* (Last accessed 16-10-2022).
7. Webpage: *https://click.linksynergy.com/deeplink?id=6m9Z7SVq*Tw&mid=42536&murl =https%3A%2F%2Fwww.edureka.co%2Fbig-data-hadoop-training-certification%3FqId%3D2eb794c4c4a8faeee4c71a45b0c54d75%26index_nam e%3Dprod_search_results_courses%26objId%3D888%26objPos%3D1&LSN SUBSITE=LSNSUBSITE* (Last accessed 16-10-2022).
8. Webpage: *https://www.awin1.com/cread.php?awinmid=6798&awinaffid=778131&clickr ef=&ued=https%3A%2F%2Fwww.edx.org%2Fcourse%2Fknowledge-management-and-big-data-in-business* (Last accessed 16-10-2022).
9. Webpage: *https://www.experfy.com/training/courses/big-data-what-every-manager-needs-to-know* (Last accessed 16-10-2022).

10. Webpage: *https://intellipaat.com/post-graduate-certification-big-data-analytics/* (Last accessed 16-10-2022).
11. Webpage: *https://linkedin-learning.pxf.io/jWWWvP* (Last accessed 16-10-2022).
12. Webpage: *https://mindmajix.com/big-data-on-aws-training* (Last accessed 16-10-2022).
13. Webpage: *http://pluralsight.pxf.io/KRMv9* (Last accessed 16-10-2022).
14. Webpage: *https://www.simplilearn.com/big-data-engineer-masters-certification-training-course?referrer=search&tag=big%20data* (Last accessed 16-10-2022).
15. Webpage: *https://click.linksynergy.com/deeplink?id=6m9Z7SVq*Tw&mid=39197&murl=https%3A%2F%2Fwww.udemy.com%2Fcourse%2Fspark-and-python-for-big-data-with-pyspark%2F* (Last accessed 16-10-2022).
16. Webpage: *https://imp.i115008.net/n111Gx* (Last accessed 16-10-2022).