DOI: https://doi.org/10.53052/9788367652018.07

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]

OKREŚLANIE KOMPETENCJI ABSOLWENTA W DATA SCIENCE: BADANIA PRACODAWCÓW

Streszczenie: Badanie zostało przeprowadzone na podstawie ankiety pracodawców opartej na ich potrzebach w zakresie kompetencji absolwentów w zakresie Data Science and Engineering, przetwarzania Big Data oraz analityki. Badanie zostało przeprowadzone w okresie od 1 września 2020 r. do 28 lutego 2021 r. Pytania zawarte w niniejszym opracowaniu miały na celu zbadanie potrzeb i oczekiwań organizacji biznesowych oraz stworzenie profili specjalistów z zakresu Big Data. Pomimo bezpośredniej ankiety pracodawców, badanie to pozwala uzyskać dodatkowe informacje na temat konkretnych umiejętności potrzebnych do profilu eksperta Big Data i specyficznych wymagań pracy Big Data.

Słowa kluczowe: Big Data, Data Science, pracodawca, profil eksperta, iBIGworld

SPECIFYING GRADUATE COMPETENCIES IN DATA SCIENCE: RESEARCH OF EMPLOYERS

Summary: The research was based on the questionnaire of the employers based on their needs in graduate competencies in Data Science and Engineering, Big Data processing, and analytics. The survey was performed during a period from the 1st of September 2020 to the 28th of February 2021. The questions in this study aimed to research the needs and expectations of business organizations and to create profiles of specialists in the field of Big Data. Despite direct employers survey this survey allows obtaining additional information on specific skills needed for a Big Data expert's profile and specific Big Data job requirements.

Keywords: Big Data, Data Science, employer, expert profile, iBIGworld

[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 cutting edge of the innovative technologies is related to the Big Data technology. With the purpose of the development of competencies the corresponding courses have been developed [1-4]. On the other hand we need to summarize the existing experience in teaching Big Data [5-16] with the requirements of labor market.
This given 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 employers based on their needs in graduate competencies in Data Science and Engineering, Big Data processing and analytics, etc. The survey was performed on-line using google forms tools. Due to various formats and specification of portals in each case the data was collected by employers and organizations that work or plan to be engaged in Big Data and Data Science domains. The survey was performed during a period from the 1st of September 2020 to 28th of February 2021. To obtain a wide range of data multiple question fields, with additional open field option, were offered to mitigate the effect of narrowed answers suggestions. The survey contains both open and close 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 organizations were selected based on their software product portfolio and engagement in Big Data filed. The survey data was presented in quantified form and statistically analysed where appropriate. The open description and free text answers were presented and analysed without modification.

Target

This survey 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" (iBIGworld) under the Erasmus+ program. The questions in this study aim to research the needs and expectations of business organizations and to create profile of specialist in the field of Big Data. Despite direct employers survey this survey allows to obtain additional information on specific skills needed for Big Data experts profile and specific Big Data job requirements.

2. Collection and analysis of data

The data was acquired by 38 organizations from four partner countries. In this research the organizations need to specify the required competencies in Big Data of graduates that need be employed. In total 38 organizations responded to the questionnaires and 38 questionaries were collected accordingly.
Sections in this work corresponds to the sections of the survey:

1. General information about organizations.
2. Big Data Job Research.
3. Broader educational competencies in Information Systems area and Big Data.
4. Competency Areas and Competency Categories.
5. Your opinion about the current level of preparation of Big Data specialists in local universities.

6. Evaluation of the competencies in terms of their importance for the employees of your organization.

3. Results

Further we study the general information about organizations.

3.1. The name of the organization/company

The employers filled the name of their organization/company. Some organizations are listed more than once (e.g., Bulgarian Post - 4 times, Escom Bulgaria Ltd – 3 times, Expert Vision Ltd - 2 times). The organization listed at No. 18 did not supply its name:

1. NAVTECH EOOD
2. Bulgarian Posts
3. Bulgarian Posts
4. Bulgarian Post
5. SoftServe
6. KEO-Zenit
7. BREN
8. Bulgarian Posts Plc
9. Expert Vision Ltd
10. IBM Bulgaria
11. Expert Vision Ltd.
12. LiLuzeNet Ltd
13. Upmind Services Co
14. Escom Bulgaria Ltd.
15. Еском България ЕООД
16. Escom Bulgaria LTD
17. ZigZag Global Ltd
18. N/A
19. Information Systems ltd.
20. ADASTRA
21. Ministry of Forejgn Affairs
22. Fijak Logic
23. Snap
24. Omega Development
25. Selleo
26. Danube Hill IT consulting
27. NIRI 4NL research and development ltd
28. Precisely
29. ROMICO GmbH
30. Nissatech
31. Rekord SI Sp. z o.o.
32. Syrmia Way
33. Code3Profit LLC
34. Way Soft
35. SkySoft
36. ING Tech Poland
37. ING Tech
38. Nignite

3.2. Indicating the type of your organization

The second question indicates the type of the organization that filled the questionnaire.

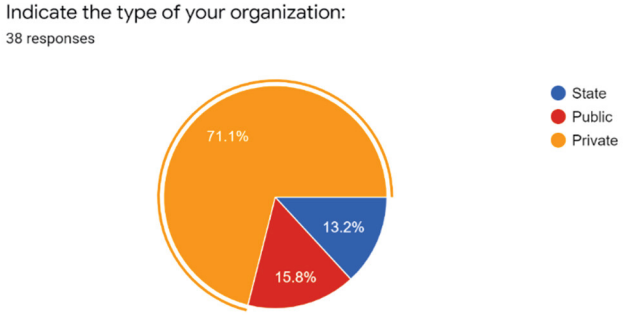

Figure 1. The type of organizations

Data description

Among 38 organizations that participated in the survey, 71.1% are private, 15.8 are public and 13.2% are state owned (Fig. 1).

Discussion

The survey shows that most of the organizations are private ones which could devise conclusion that they are more agile in adopting new Big Data technologies and methods and thus, requiring more graduates with specific DS and BG skills and expertise.

3.3. Size of the company - staff number

The question asks for the size of the company measured the by the staff number.

Data description

Among 38 organiyations, 44.7% of them have more than 50 empleyees, 26.3% of orgqanizations have beetwenn 10 and 50 employees, and the rest of the, 28.9 comapnies are the small ones having less than 10 employees (Fig. 2).

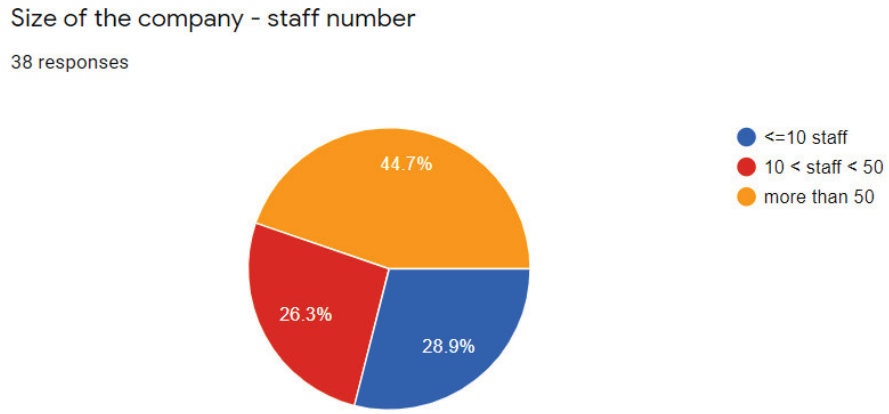

Figure 2. Distribution of the organization's size

Discussion

The question shows that medium and large organizations are eager to work in a new and challenge areas, like BD and DS, that impose steep leraning curve and require specific graduates comppoetences.

3.4. What segment of IT of the industry do you work in?

This question aims to find the segment of IT industry in whuich these organisations mainly work, among offered IT industry domains.

Data of organizations

Most of the companies (39.5%) consider themselves working in IT services sector, the next group includes organizations working in IT outsourcing (15.8%), then two group of organizations working in IT outstaffing and other product startups and the rest of them belongs to other domains (Fig. 3).

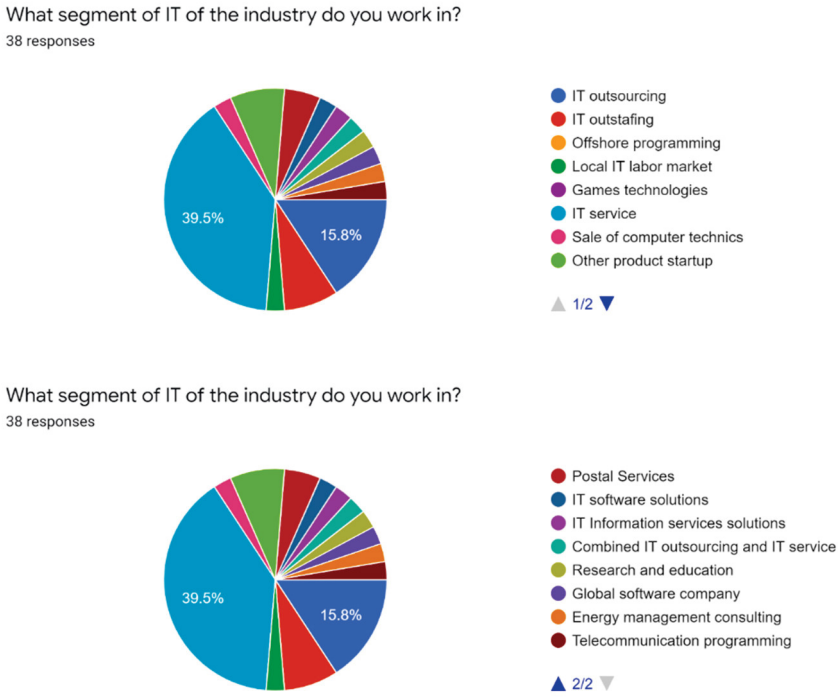

Figure 3. The IT segments of the companies.

Discussion

The organizations come from various IT sectors but most of them work in the IT services and outsourcing.

3.5. Indicating the main direction(-s) of activities of your organization

The question aims to collect the main direction(-s) of business and IT development activities of the organizations, allowing multiple choices of the domains.

Data description

Most of the companies (50%) mentioned Manufacturing and development as the main direction of their activities. The number of the organizations work in the Consulting domain (12 – 31.6%) as well, 10 organizations indicate that their main activity is also related to Design, the same number mentioned Sales as their activity, and significant

numbers of organizations indicate Research (9 – 23.7%), Public services (7 – 18.4%) and Customer service (7 – 18.4%) (Fig. 4).

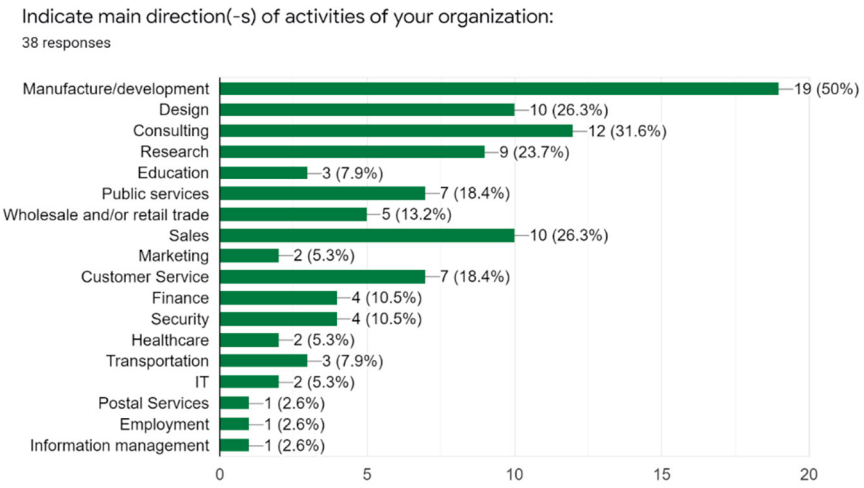

Figure 4. Main activities

Discussion

The main directions of the organizations answering this question are in manufacturing, design, development and consulting which shows their interest in improving their business toward innovative Big Data and Data science/engineering -based products and services.

3.6. Do you use Big Data in your business?

The question aims to find whether organizations use Big data in their business activities.

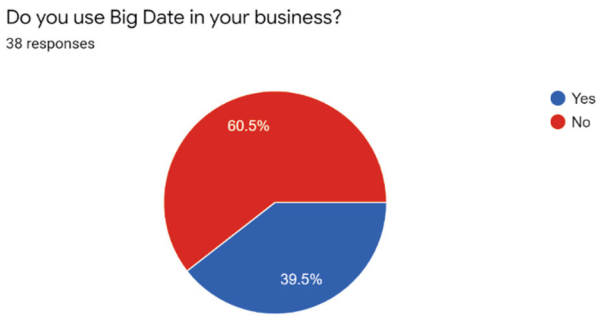

Figure 5. Usage of Big Data

Data description

Among 38 organizations, the majority of them (60.5%) do not use Big Data, but intend to do so, while the rest 39.5% use Big Data in their business activities (Fig. 5).

Discussion

More than a half of the organizations have not used Big Data in their business yet, but participating in this survey clearly express their interest to improve their business toward Big Data – related domains and activities.

The following research was dealing with Big Data jobs.

3.7. What jobs do you offer?

Since only some organizations (15 of 38) either use or plan to use Big Data the questions aim to find out what kind of job related to Big Data they currently offer.

Data description

Most of the organizations (14 – 93.3%) offer Data Analyst job, Data Scientist job is offered by 11 of them (73.3%), Data Architect job is offered by 10 organizations (66.7%), Database Manager job is offered by 9 organizations (60%), and Big Data Scientist and Security Engineer jobs are offered by 7 organizations (46.7%) (Fig. 6).

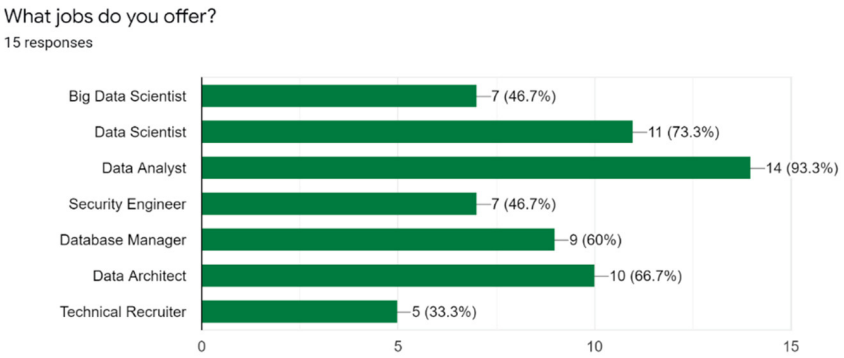

Figure 6. Job offers

Discussion

Less that a half of the organizations responded to this question (~40%) showing that most of them have not still recognized the need for Big Data and Data Science -related specialists, but it is expected that they will move their strategy to this domain, as soon as there are more BD/DS graduates and job offers in these areas.

3.8. Does an experience in a Big Data field is required?

This question aims to discover if the organizations need an experience in Big Data field from their employees. Only 15 organizations respond to it.

Data description

Most of the organizations require short practice in Big Data field, up to one year (46.7%) and almost equal number of them requires experience from 1 to 3 years (40%). The same number of organizations (exact one) require either significant experience (4-5 years) or no experience at all (Fig. 7).

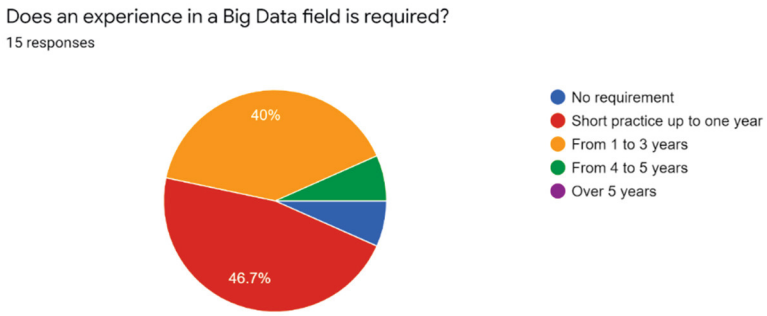

Figure 7. The requirement of the experience

Discussion

The organizations prefer at least one year of experience and some of them require up to three years and more which shows that there is a need for a master degree on Big Data and Data Science offered at the Universities or shorter specialization and focused courses delivered by both academic and R&D orgnizations.

3.9. Rating the competences of academic/analytical Big Data employees

The employers need to rate the competences of academic/analytical Big Data employees they require in their activities.

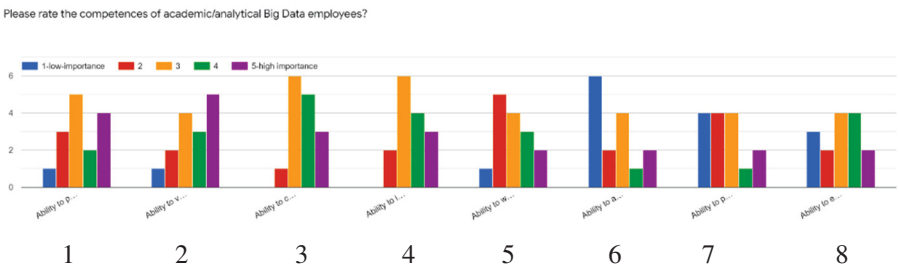

Figure 8. Rating the "academic" competencies

Data description

The organizations rate the following competences they need from their employees with respect to Big Data expertise and mark them from low importance (1) to high importance (5) (Fig. 8):

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.

Discussion

Aside for being neutral, most organizations prefer abilities to perform simulations, experiments and verification with statistical and BD/DS tools, and work on new technologies and methods. Skills of writing technical/research papers, application of modern psychology and pedagogy and ability of patent inventions are of low priority for most of the organizations.

3.10. What data competencies are needed?

The question needs to point out the competences needed in data management and analytics (Fig. 9).

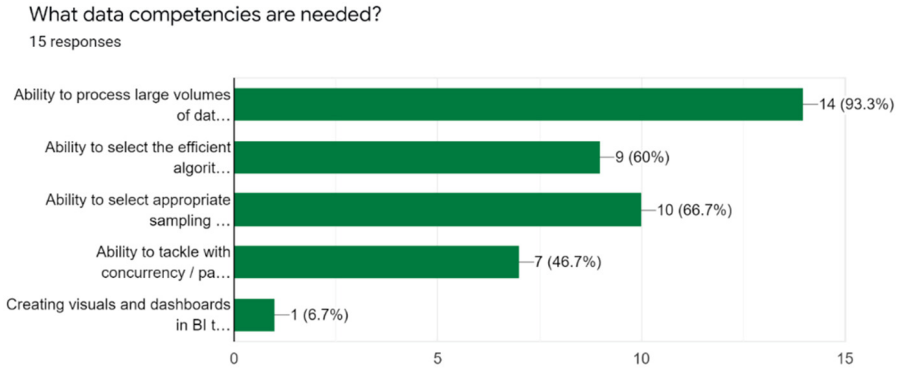

Figure 9. Data competencies

Data description

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.
5. Creating visuals and dashboards in BI tools (e.g. Looker, Tableau, Power BI, Google Data Studio), Experience with analytical packages such as SciPy or

R, Understanding of Snapchat with great product sense and product understanding, Ability to communicate the results of analyses in a clear and effective manner to a senior audience.

Discussion

Most organizations express the need for abilities to process large volumes of data using hierarchical storage, hashing and filtering and to select the efficient algorithm to Big Data, and to select appropriate sampling and filtering method for given Big Data analysed case.

3.11. What related data competencies are needed?

This question aims to define related data competences needed by organizations (Fig. 10).

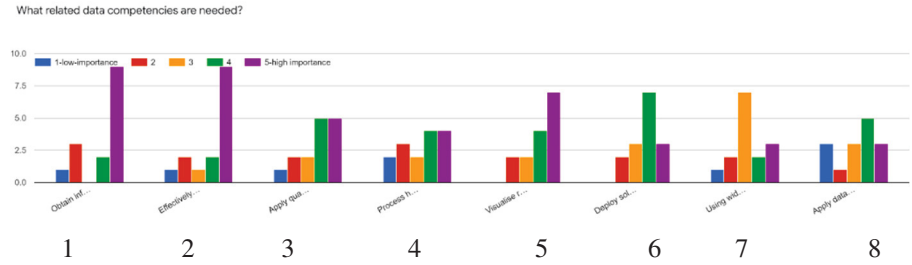

Figure 10. Related data competencies

Data description

The organizations rate the following additional data related competences they need from their employees and mark their importance from low (1) to high (5):

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. Apply data security mechanisms and controls at each stage of the data processing.

3.12. Selecting the Big Data tools that employee should know

This question aims to identify Big Data tools that employee should know to be admitted to the 15 organizations that offer the jobs in Big Data domain (Fig. 11).

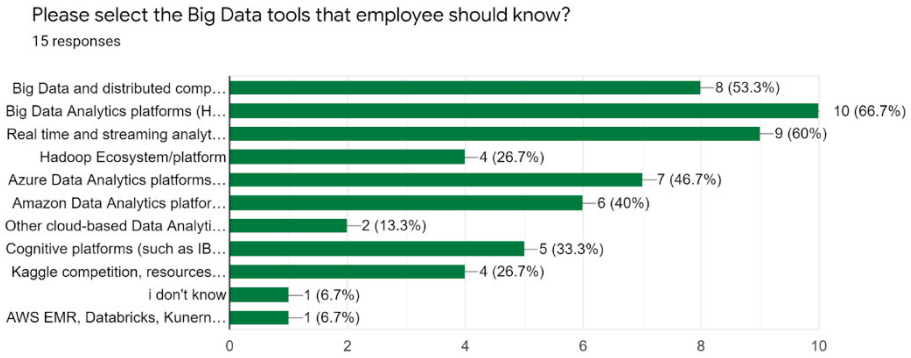

Figure 11. Big Data tools required

Data description

The organizations identify skills and expertise in Big Data that they need and ask in jobs offered.

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.
10. AWS EMR, Databricks, Kubernetes.

Discussion

Most of the organizations look for skills and expertise in Big Data processing and analytics tools and platforms, for offline and streaming data, such as Apache Spark, MapReduce/Hadoop, Lucene, NLTK, Pregel, etc. Flume, Kafka, Storm. Several organizations obviously base their Big Data business activities on cloud service providers, such as Amazon and Microsoft Azure, and ask for skills and knowledge of their Big Data analytics services. Other cloud-based data analytics platforms, such as Hortonworks, Vertica LexisNexis HPCC System, and cognitive platforms, such as IBM Watson, Microsoft Cortana, etc. are less required.

3.13. What programming languages are required?

The question aims to discover what programming languages are required for the potential employees (Fig. 12).

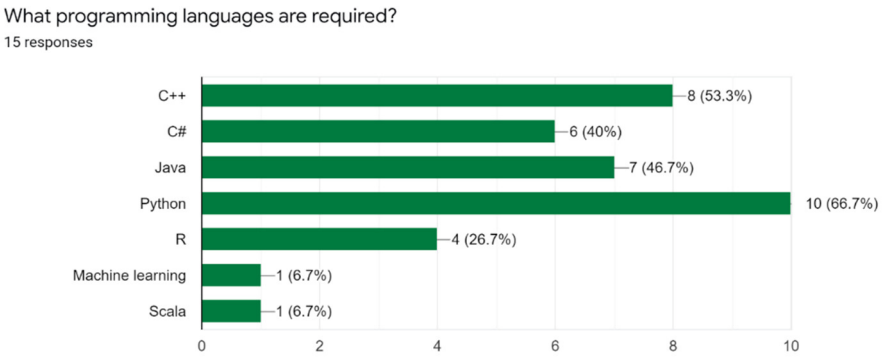

Figure 12. Programming languages required

Data description

Among 15 responses, most of the organizations that respond require Python programming language (10-66,7%), then C++ (8 – 53.3%), Java (7 – 46,7%), C# (6 – 40%) and R (4 – 26.7%).

Discussion

As expected, most of the companies require the knowledge of Python programming language as the primary one for the Data Science and Analytics, while also prefer other well-established general-purpose programming languages, C/C++, Java and C#.

3.14. What additional job competencies are required?

The organizations were asked to specify additional competence they require from the potential employees in a free form.

Data description

The answers to this question specify the competences from technial and formal ones, such as SQL, Linux, Kerberos, to more general and non-formal, such as soft skills and ability to colaborate with engieers. The competeneces required listed in original form are:
Computer Literacy, Technical background in Servers and Storages, Soft skills, domain knowledge, Fluent English (communication, presentation in English and writing), Experience with SQL and building data pipelines, Strong statistical knowledge, Ability to collaborate with engineers, product managers, designers and other cross-functional teams, An understanding of Snapchat with great product sense and product understanding, Excellent interpersonal skills, including verbal and written communication, Ability to initiate and drive projects to completion with minimal guidance experience, Cloud technologies, Business analysis, Openmind, Understaning web apps developlemnt, Understanding of data characteristics and methodical approach to data analytics, SQL, SAS, Financial reporting, Bash, Linux, Groovy, Scala, ETL, Kerberos, SRUM/Agile, distributed systems, business domain knowledge.

Discussion

The organizations require additional skills such as so-called soft skills and fluent English, as expected and competences related to Big Data and Data Science domains, such as Cloud technologies and Business analytics tools and platforms.

3.15. Your impressions like an employer for the specialists, graduating with a training degree in Big Data.

By answering to this question, the organizations need to express their impression regarding the level of knowledge and expertise of the students graduated in Big Data domain.

Data description

The impression of the organization regarding the competences of the Big Data graduates are specified in a free form, as follows:

- Additional knowledge and practice are needed.
- I don't have much experience.
- Usually graduates have good basic knowledge in programming and databases, but lack experience with up-to-date data analytics platforms and tools, and proper statistics skills.
- Cool.
- Big Data education for graduates is important aspect to provide good pool of talents.
- Problem solving knowledge is missing.
- Limited knowledge what tools and techniques are currently available.
- Our suggestions: open minded, team player, good organization, improve technical skills, reporting and analysing.
- Positive value of graduates.

Discussion

The employers estimate the skills and expertise in Big Data and Data Science domains after graduating and some of them pointed out that additional knowledge an practice are needed, as well as more experience in BD problem solving which justify the objectives of the iBIGworld project and will give excellent applications of project's results.

The following study is related to the broader educational competencies in Information Systems area and Big Data. Namely, the impressions like an employer for the specialists, graduating with a training degree in an information system are investigated.

3.16. Indicating how satisfied you are as an employer with the level of preparation of Master's studies' graduates (or high school graduates in general) in the area of information systems and Big Data

The question aims to indicate the level of satisfaction of employers with the expertise of the Master students in the field of Big Data (Fig. 13).

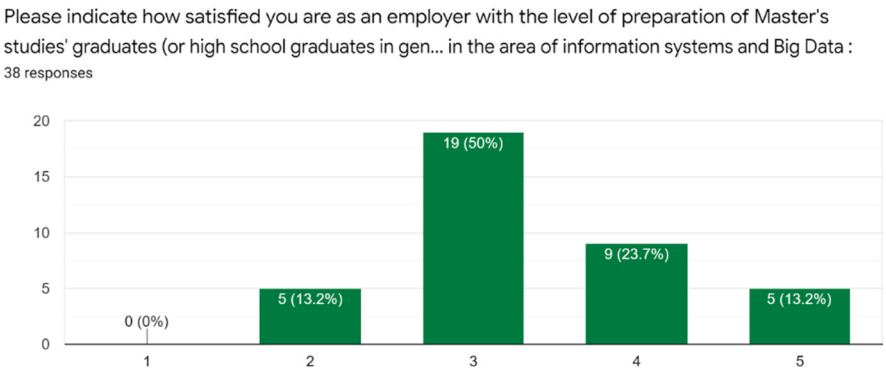

Figure 13. Satisfaction with the level of preparation

Data description

Among 38 responses from all organizations that participated in the survey a half of them (19 – 50%) are in the middle of the satisfaction (neutral), having 9 organizations (23.7%) satisfied and 5 organizations (13.2%) very satisfied. Only 5 organizations (13.2%) are slightly not satisfied.

Discussion

The fact that more than a half of the organizations are not quite satisfied by the level of preparation of Masters' graduates in IS and Big Data domains, gives the opportunity and a "fertile ground" for the project goals and results to improve their satisfaction offering appropriate training and teaching in BD and DS domains.

3.17. Satisfaction with the supply (quantity) of information systems specialists available for hire on the job market today and can they work with Big Data

The question has itention to examine the satisfaction of the organizations by the quantity of IT and IS specialists available for hirining on the job market that could sucessfully work in the field of Big Data (Fig.14).

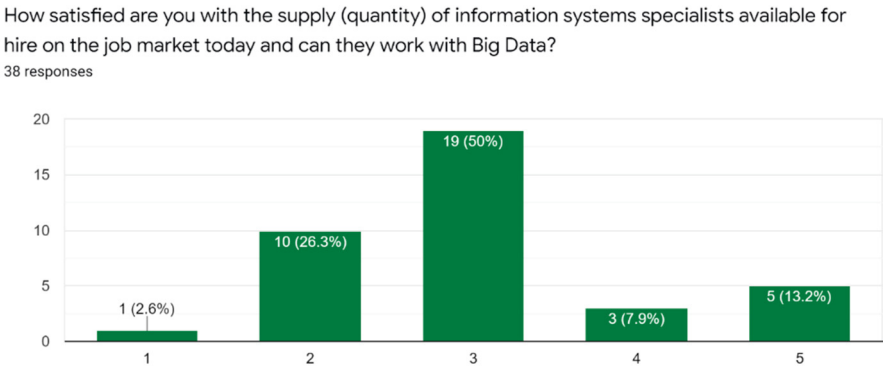

Figure 14. Satisfaction with the quantity of IT specialists

Data description

Among 38 responses from all organizations that participated in the survey a half of them (19 – 50%) are in the middle of satisfaction, having 3 organizations (7.9%) satisfied and 5 organizations (13.2%) very satisfied. But 10 organizations (26.3%) are slightly not satisfied, and one organization is not satisfied at all.

Discussion

More than 80% of the organizations are note quite satisfied with the supply of BD and DS specialists available on the job market and their skills in Big Data, so the objectives and results of the project need to improve this situation offering more graduates with specific skills and expertise in BD and DS tools. Technologies and platforms ready for jobs offered in the market.

3.18. Views towards raising the qualification of current employees of your organization by letting them study Big Data at a master's level

The organizations answer this question in order to indicate their view towards raising the qualification of current emploeyees by their study Big Data at a master's level (Fig.15).

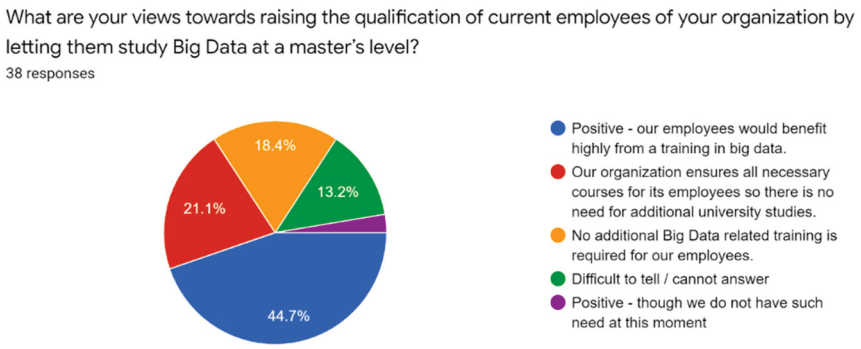

Figure 15. Assessment of raising the qualifications

Data description

Almost half of the organizations (44.7%) consider such approach as positive stating that their employees would highly benefit from training in Big Data. Some of the organizations (21.1%) provides necessary courses for their employees in this field without the need for additional master studies. The rest of the organizations either do not need additional Big Data related training for their employees (18.4%) or could not detect the need (13.2%), or do not have such need at the moment.

Discussion

The answers to this question show that there is a great need for Big Data and Data Science related study at the master level, but also appreciate the specific courses

offered by the companies and organizations themselves to their employees, which both ar4 in line with the project goals.

The following questions are related to the competency areas and competency categories. In this section organizations can specify the competencies that graduates of Big Data need to have, a what kind of Big Data specialists they actually need regarding skills and expertise in various areas, such as:

- Business Continuity and Information Assurance.
- Systems Development and Deployment.
- Data, Information and Content Management.
- Enterprise Architecture.
- IS Management and Operations.
- IT Infrastructure.

3.19. Business Continuity and Information Assurance

In this question organizations can specify the competencies that graduates of Big Data need to have, a what kind of Big Data specialists they actually need regarding various fields within Business Continuity and Information Assurance domain (Fig. 16).

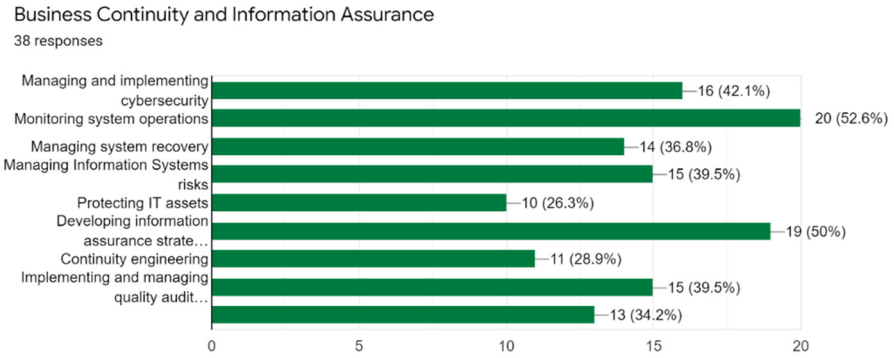

Figure 16. Required competencies

Data description

1. Managing and implementing cybersecurity.
2. Monitoring system operations.
3. Managing system recovery.
4. Managing Information Systems risks.
5. Protecting IT assets.
6. Developing information assurance strategy.
7. Continuity engineering.
8. Implementing and managing quality audit processes.
9. Assuring safety throughout systems lifecycle.

Discussion

The answers to this question emphasize the need for competences in monitoring system operations and developing information assurance strategy, but also express the importance of managing an implementing cybersecurity and implementation and management of the quality audit process. Mentioned competences should be used in account when preparing Big Data courses within the project aside from BD and DS skills and competencies.

3.20. Systems Development and Deployment

In this question organizations can specify the competencies that graduates of Big Data need to have, a what kind of Big Data specialists they actually need regarding various fields within Systems Development and Deployment domain (Fig. 17).

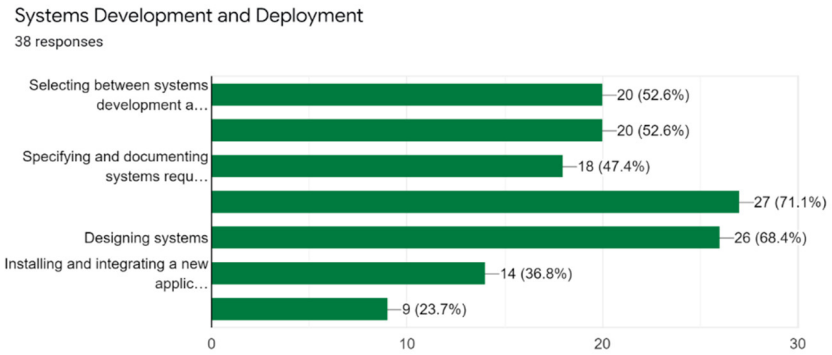

Figure 17. Required competencies

Data description

1. Selecting between systems development approaches.
2. Managing plan-based, hybrid, and agile development approaches.
3. Specifying and documenting systems requirements.
4. Implementing and testing an application.
5. Designing systems.
6. Installing and integrating a new application.
7. Managing external systems development resources.

Discussion

The answers to this question emphasize the need for competences in designing, implementing and testing BD applications, as expected, but also ability to select and apply various system development approaches and managing agile practices in software development. Mentioned competences should be used in account when preparing Big Data courses within the project aside from BD and DS skills and competencies.

3.21. Data, Information and Content Management

In this question organizations can specify the competencies that graduates of Big Data need to have, a what kind of Big Data specialists they actually need regarding various fields within Data, Information and Content Management domain (Fig. 18).

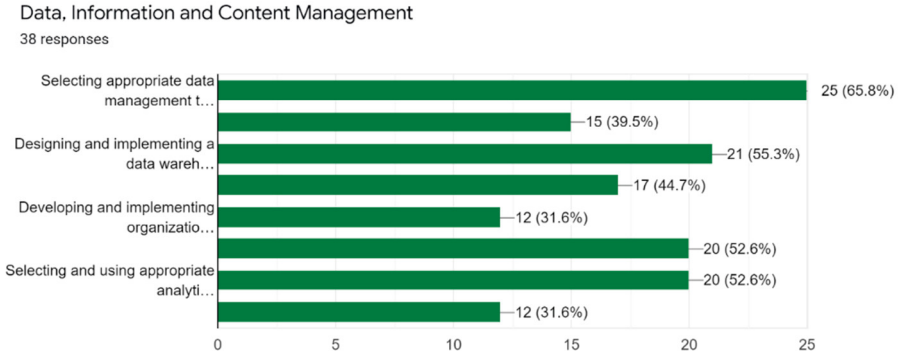

Figure 18. Requirements related to the data management

Data description

1. Selecting appropriate data management technologies based on the needs of the domain.
2. Securing domain data and protecting user privacy and organizational intellectual property using appropriate technical solutions.
3. Designing and implementing a data warehouse using a contemporary architectural solution.
4. Creating a scalable infrastructure for large amounts of data using parallel and distributed technologies.
5. Developing and implementing organizational information management policies and processes.
6. Integrating and preparing data captured from various sources for analytical use.
7. Selecting and using appropriate analytics methods.
8. Designing and implementing architectures for organizational content management systems.

Discussion

The answers to this question put the focus on competences related to selection of appropriate data management technologies based on the needs of the BD domain, selection and usage of appropriate analytical methods and their integration in analytics of Big Data collected from various sources. The necessary competences also include design and implementation of data warehouse solutions. Mentioned competences should be used in account when preparing Big Data courses within the project aside from BD and DS skills and competencies.

3.22. Enterprise Architecture

In this question organizations can specify the competencies that graduates of Big Data need to have, a what kind of Big Data specialists they actually need regarding various fields within Enterprise Architecture domain (Fig. 19).

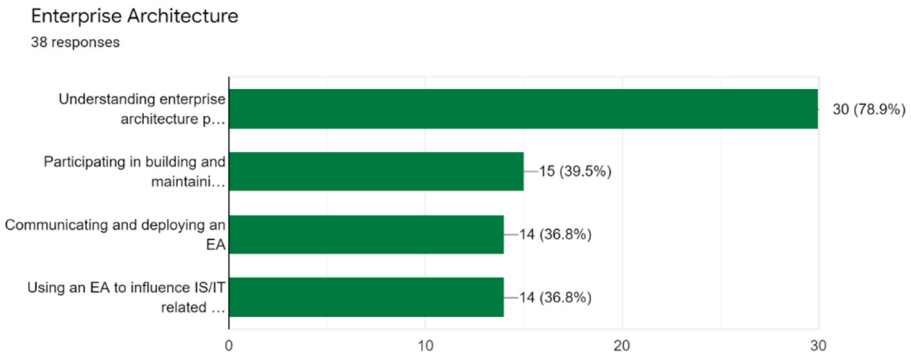

Figure 19. Competencies on enterprise architecture

Data description

1. Understanding enterprise architecture principles and the value it provides to businesses.
2. Participating in building and maintaining an EA.
3. Communicating and deploying an EA.
4. Using an EA to influence IS/IT related organizational improvement projects.

Discussion

The answers to this question emphasize the highest need for competences in understanding enterprise architecture principles and the value it provides to business activities and processes related to Big Data processing and analytics. The importance of building, deploying and maintainig Big Data – based enterprise architecture are equally specified. Mentioned competences should be used in account when preparing Big Data courses within the project aside from BD and DS skills and competencies.

3.23. IS Management and Operations

In this question organizations can specify the competencies that graduates of Big Data need to have, a what kind of Big Data specialists they actually need regarding various fields within IS Management and Operations domain (Fig. 20).

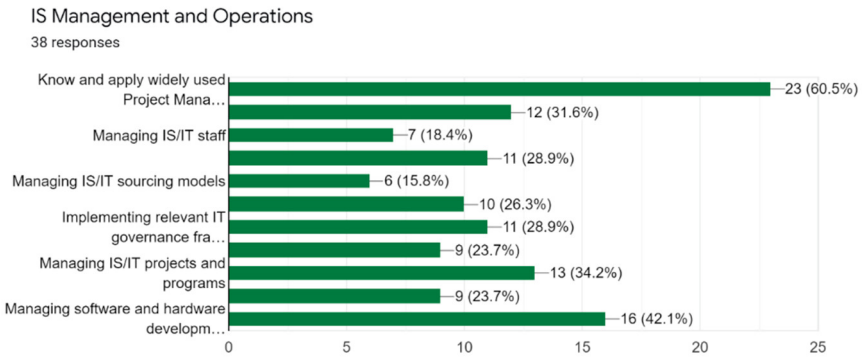

Figure 20. Requirements on IS management

Data description

1. Know and apply widely used Project Management tools and techniques.
2. Managing the IS/IT function.
3. Managing IS/IT staff.
4. Managing IS/IT service production.
5. Managing IS/IT sourcing models.
6. Managing and coordinating information resources.
7. Implementing relevant IT governance frameworks withing the organization based on strategic guidance.
8. Understanding laws and regulations directly affecting IS/IT management and operations.
9. Managing IS/IT projects and programs.
10. Managing IS/IT project portfolios.
11. Managing software and hardware development and maintenance.

Discussion

The answers to this question put the focus on competences related to selection of appropriate data management technologies based on the needs of the BD domain, selection and usage of appropriate analytical methods and their integration in analytics of Big Data collected from various sources. The necessary competences also include design and implementation of data warehouse solutions. Mentioned competences should be used in account when preparing Big Data courses within the project aside from BD and DS skills and competencies.

3.24. IT Infrastructure

In this question organizations can specify the competencies that graduates of Big Data need to have, a what kind of Big Data specialists they actually need regarding various fields within IT Infrastructure domain (Fig. 21).

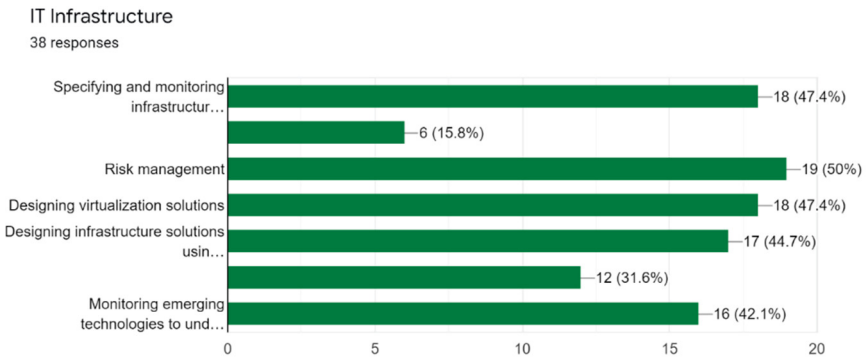

Figure 21. Competencies on IT infrastructure

Data description

1. Specifying and monitoring infrastructure contracts.
2. Negotiating contracts and managing infrastructure vendors.
3. Risk management.
4. Designing virtualization solutions.
5. Designing infrastructure solutions using external service provider(s) (cloud computing).
6. Maintaining a set of standards and policies and understand the key laws and regulations to relevant infrastructure decisions.
7. Monitoring emerging technologies to understand their potential to support the domain.

Discussion

The answers to this question almost equally prioritize competences related to IT infrastructure in development, deployment and maintenance of BD solutions, such as specifying and monitoring infrastructure contracts, risk management and developing and operation of virtualization solutions using external cloud service provider(s). Mentioned competences should be used in account when preparing Big Data courses within the project aside from BD and DS skills and competencies.

The next study is related to the opinion about the current level of preparation of Big Data specialists in local universities. In this section the organizations need to evaluate the level of knowledge and expertise in Big Data that diploma/master engineers gather from local universities.

3.25. Indicating the level of agreement with the following statements about the university graduates in the area of information system (or Big Data).

This question aims to collect the level of agreement with the statements regarding preparedness od university graduates for Big Data area, from strongly disagree (1) to strongly agree (5) (Fig. 22).

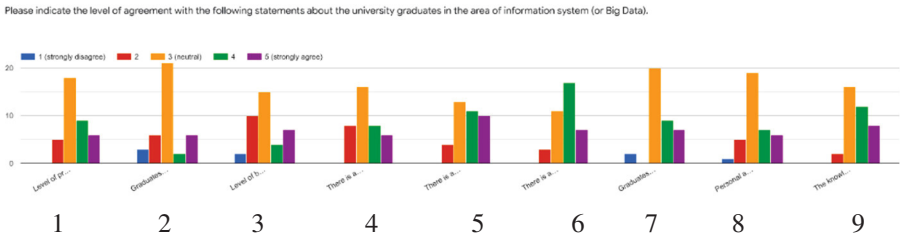

Figure 22. Level of the graduates

Data description

1. Level of professional theoretical knowledge is high.
2. Graduates are well trained practically, they know how to apply theoretical knowledge in practice.
3. Level of basic knowledge in such areas as business management, economics, and law is good.
4. There is a good understanding and knowledge about the latest IS related international standards.
5. There is a strong will to create and innovate.
6. There is a strong will and commitment to finding and keeping a good job.
7. Graduates respect the work ethics of the corporate culture.
8. Personal ambitions are measured and reasonable, self-assessment – realistic.
9. The knowledge of English (or other relevant foreign language) is good.

Discussion

The employers are manly neutral to all statements, disagreeing on the statement related to high level of professional theoretical knowledge and that graduates are well trained practically, and that they know how to apply theoretical knowledge in practice. The employers agree more on statements related to strong will to create and innovate, to find and keep a good job and that the knowledge of English (or other relevant foreign language) is good.

3.26. Rating the idea of building a website which will present results Big Data research carried out by local University employees

The organizations need to rate the idea of building a Web site that will contain the results of the Big Data research carried out by Universities (Fig. 23).

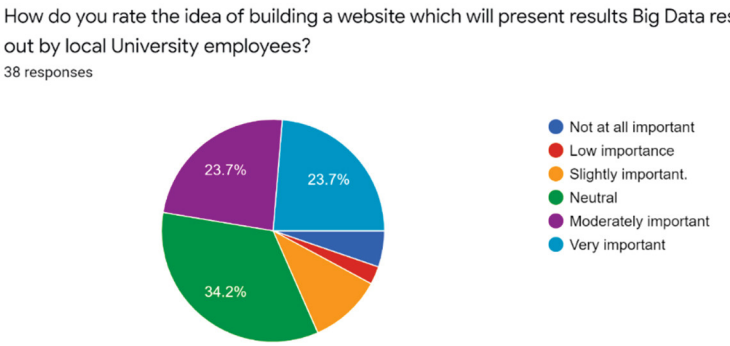

Figure 23. Building the website

Data description

Most of the organizations rate this idea as Very important (23,7%) and Moderately important (23,7%), which are almost half of them, but a significant number are neutral to this idea (34,2%), while some small number of organizations of the rate it as low importance or not important at all.

Discussion

The answers to this question emphasize the need to build a website which will present results of Big Data research and development carried out by local Universities, so such initiative will be among the project goals.

The next study is dealing with the evaluation of the competencies in terms of their importance for the employees of your organization. In this section the organizations need to evaluate the various competences with regard to their importance for their employees.

3.27. Evaluating the following competencies in terms of their importance for the employees of the organization that are working in positions requiring high level education (master's or equivalent) in Information Systems or related areas - Systemic competences.

This question aims to evaluate the necessary **Systemic competences** with regard to their importance for the employees having master degree or equivalent in Computer Science, Information Systems, Data engineering, Data Science, etc. from strongly disagree (1) to strongly agree (5) (Fig. 24).

Figure 24. Systemic competencies

Data description

1. Ability to develop intellectually and culturally.
2. Ability to specify the essence of problems in the professional domain, and to find adequate solutions.
3. Ability to generate new ideas and come up with out-of-the-box ways for their implementation.
4. Ability to make management decisions, evaluate their possible effects, and embrace the responsibility for the work results of your team.
5. Ability to learn new methods and to change the nature of his/her job.
6. Ability to head projects, organize teamwork, and to initiate positive business process changes.
7. Ability to carry out research including analysis of problems, formulation of aims and tasks, selection of research methods, as well as evaluation of quality.

Discussion

The answers to this question point out the importance (medium or high) of all mentioned systemic competence related, so building such competence must be taken in account when achieving the project's goals.

3.28. Evaluating the following competencies in terms of their importance for the employees of your organization that are working in positions requiring high level education (master's or equivalent) in Information Systems or related areas (2/4) - Methodological competences

This question aims to evaluate the necessary **Methodological competences** with regard to their importance for the employees having master degree or equivalent in Computer Science, Information Systems, Data engineering, Data Science, etc. from strongly disagree (1) to strongly agree (5) (Fig. 25).

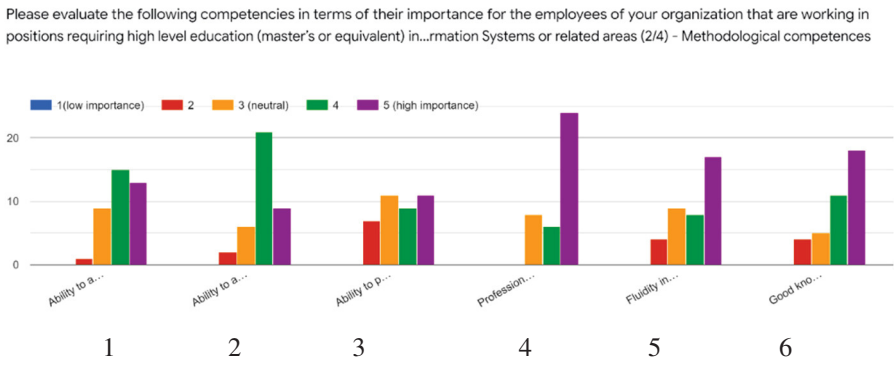

Figure 25. Methodological competencies

Data description

1. Ability to analyze problems using system analysis, analysis by synthesis, and other methods.
2. Ability to analyze, verify, evaluate completeness of data, and to supplement or create it under the conditions of uncertainty.
3. Ability to propose concepts, models, create and adapt methods and tools for professional activities using resources from natural, social, economic sciences, and humanities.
4. Professional computer and IT usage skills.
5. Fluidity in spoken and written native language.
6. Good knowledge of English (or other relevant foreign language).

Discussion

The answers to this question point out the importance (medium or high) of all mentioned methodological competence especially those related to professional computer and IT usage skills, fluidity in spoken and written native language as well as good knowledge of English (or other relevant foreign language). Building such competence should be taken in account when achieving the project's goals.

3.29. Evaluating the following competencies in terms of their importance for the employees of your organization that are working in positions requiring high level education (master's or equivalent) in Information Systems or related areas (2/4) - Social/personal competences

This question aims to evaluate the necessary **Social/personal competences** with regard to their importance for the employees having master degree or equivalent in Computer Science, Information Systems, Data engineering, Data Science, etc. from strongly disagree (1) to strongly agree (5) (Fig. 26).

Figure 26. Social/personal competencies

Data description

1. Ability to follow principles of social responsibility as well as legal and ethical norms when developing own career or building a new business.
2. Ability to organize and manage communication (also multi-cultural).
3. Ability to work professionally in an international environment that involves a lot of mobility.
4. Ability to define and communicate objectives of professional activities.
5. Ability to lead by example and inspire others.
6. Ability to solve social and personal problems as well as problems resulting from different worldviews and cultures.
7. Ability to lead teams consisting of people with different levels of skills and varying specializations.
8. Ability to work efficiently in national and international teams.
9. Ability to work individually with minimum guidance, manage one's work and time.

Discussion

The answers to this question point out the importance (medium or high) of all mentioned social/personal competences competence, so building such competence should be included in the project's results.

3.30. Evaluating the following competencies in terms of their importance for the employees of your organization that are working in positions requiring high level education (master's or equivalent) in Information Systems or related areas (3/4) - Professional competences: analysis, design, and project management

This question aims to evaluate the necessary **Professional competencies** (**analysis, design, and project management**) with regard to their importance for the employees having master degree or equivalent in Computer Science, Information Systems, Data engineering, Data Science, etc. from strongly disagree (1) to strongly agree (5) (Fig. 27).

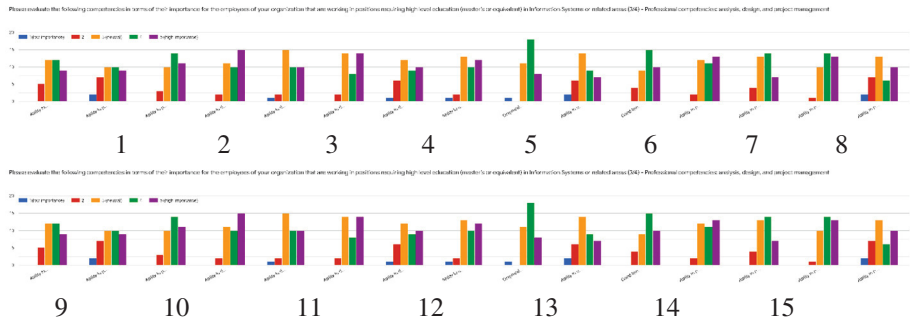

Figure 27. Professional competencies

Data description

1. Ability to model, analyze, and evaluate organization's business processes.
2. Ability to perform semantic modeling as well as conceptual modeling of various domains and information systems.
3. Ability to perform information systems requirements analysis and specification.
4. Ability to design relational databases.
5. Ability to design NoSQL (or other non-relational) databases.
6. Ability to design local, network-based, and web information systems.
7. Ability to design grid-based systems and systems driven by the principles of cloud computing.
8. Ability to carry out system reengineering and to analyze legacy databases.
9. Graphical user interface analysis and design skills.
10. Ability to use computer aided software engineering (CASE) tools.
11. Good knowledge of latest information systems modeling languages and standards.
12. Ability to prepare high quality project documentation, also to write comprehensive user manuals.
13. Ability to plan and manage information systems development projects whether using classic or agile approaches.
14. Ability to plan and carry out comprehensive analysis of datasets that are diverse, complex, stored in numerous databases, and very large (Big Data skills).
15. Ability to perform market research and write business plans on information systems, software products, or IS development tasks.

Discussion

The answers to this question point out the importance (medium or high) of all mentioned social/personal competences competence, so building such competence should be included in the project's results.

3.31. Evaluating the following competencies in terms of their importance for the employees of your organization that are working in positions requiring high level education (master's or equivalent) in Information Systems or related areas (4/4) - Professional competences: implementation and systems administration

This question aims to evaluate the necessary **Professional competencies** (**implementation and systems administration**) with regard to their importance for the employees having master degree or equivalent in Computer Science, Information Systems, Data engineering, Data Science, etc. from strongly disagree (1) to strongly agree (5) (Fig. 28).

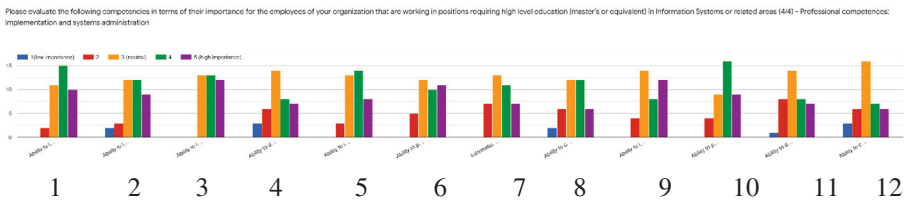

Figure 28. Implementation and systems administration

Data description

1. Ability to implement and administer relational databases.
2. Ability to implement and administer NoSQL databases.
3. Ability to implement software components of information systems following given design documents and using various programming languages and tools.
4. Ability to develop systems for the Semantic Web.
5. Ability to integrate various software and hardware components of integrated information systems locally or involving networks.
6. Ability to plan, perform, and manage manual or automated testing.
7. Information systems administration skills.
8. Ability to configure and manage comprehensive ERP (enterprise resource planning) or similar available systems.
9. Ability to implement cloud computing based solutions.
10. Ability to perform data analysis using the principles of data mining and business intelligence, ability to create data visualizations.
11. Ability to develop IS modelling and development solutions, including standards and CASE tools.
12. Ability to certify software development products.

Discussion

The answers to this question point out the importance (medium or high) of all mentioned social/personal competences competence, so building such competence should be included in the project's results.

3.32. Evaluating the following competencies in terms of their importance for the employees of your organization that are working in positions requiring high-level education (master's or equivalent) in Information Systems or related areas (4/4) - Research and academic/analytical competences

This question aims to evaluate the necessary **Research and academic/analytical competences** with regard to their importance for the employees having master degree or equivalent in Computer Science, Information Systems, Data engineering, Data Science, etc. from strongly disagree (1) to strongly agree (5) (Fig. 29).

Figure 29. Research competence

Data description

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

Discussion

The answers to this question point out the following competences and abilities of the employees having master's degree in CS or related filed: to innovate and modify methods and approaches used in the organization, to perform simulations and experiments, and to critically analyze their results using statistical methods, as well as to carry out feasibility studies on new technologies, methods, and standards that could be of use to the organization. The results that will be developed during the course of the project will focus on these skills and abilities of the master graduates to satisfy the needs and expectations of their employers.

4. Conclusions

So, the research was performed based on the questionnaire of the employers based on their needs in graduate competencies in Data Science and Engineering, Big Data processing, and analytics. The questions in this study aimed to research the needs and expectations of business organizations and to create profiles of specialists in the field of Big Data. Despite direct employers survey this survey allowed us to obtain additional information on specific skills needed for a Big Data expert's profile and specific Big Data job requirements.

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_name%3Dprod_search_results_courses%26objId%3D888%26objPos%3D1&LSNSUBSITE=LSNSUBSITE* (Last accessed 16-10-2022).
8. Webpage:*https://www.awin1.com/cread.php?awinmid=6798&awinaffid=778131&clickref=&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 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).