



**СЪВРЕМЕННИ СТРАТЕГИИ И ИНОВАЦИИ
В УПРАВЛЕНИЕ НА ЗНАНИЕТО**



**CONTEMPORARY STRATEGIES
AND INNOVATIONS
IN THE KNOWLEDGE MANAGEMENT**



**ЗА БУКВИТЕ
О ПИСМЕНОСТЪ**

С Ъ Д Ъ Р Ж А Н И Е

ПРЕДГОВОР	11
ПОЗДРАВИТЕЛНИ АДРЕСИ	19
1. Камелия Планска-Симеонова & Светослава Димитрова. Основни аспекти на законодателството в областта на авторското право в Европа и Югоизточна Азия.....	27
2. Соня Спасова. Междуинституционално сътрудничество за интегриране и социализация на културното наследство в училище	42
3. Елена Савова. Комуникативната роля на университетските библиотеки.....	48
4. Диана Стоянова & Християн Атанасов. Лозенград и региона на страниците на българската възрожденска преса (според дигитализирани периодични издания в Национална библиотека „Св. св. Кирил и Методий“)	54
5. Калина Минчева. Български авторски книги за деца от ХХ век.....	66
6. Пепа Петрова. Подход за подготовка на големи данни за анализ.....	70
7. Мириана Павлова. Импиджът и доверие в университета – ключови фактори при подсилването на мотивацията на обучаемите	77
8. Поли Муканова. Преглед на институционалните практики за четене в България в историческа перспектива	83
9. Христо Тутунаров. Блскуотър – първи от многото	94
10. Десислава Стоева. Услуги за сигурност на Република България под контрола на Парламента: условия и механизми	99
11. Боян Аспарухов. Документалното наследство на кореспондентите от Руско-турската освободителна война 1877 – 1878 – част от културното наследство в европейския югоизток.....	103
12. Ева Ценкова. Европейска подкрепа и помощ на младото поколение – познанието от университета до общинската мрежа в България	107

13. Симона Душева. Образи и цветове по долното течение на река Тунджа. Представяне на изследователски филм за обществените процеси и културно-историческото наследство в района на община Елхово	112
14. Събина Ефтимова. Учене чрез правене в съвременна обществена среда: пример от обучението на студенти в специалност Библиотечен и информационен мениджмънт в УниБИТ	114
15. Десислава Бакърджиева. Телевизията и нейното влияние върху междуличностната комуникация и появата на социални общности.....	118
16. Мирияна Павлова & Диана Стоянова & Кристина Босакова. Стратегия за създаване и управление на университетски имидж.....	124
17. Мартин Захариев. Аспекти на защита на данните посредством изкуствен интелект в търговския арбитраж.....	129
18. Ива Костадинова & Кристиян Алексиев. Оптимизация на WEB-тестови страници в система за с-тестове.....	140
19. Петя Петкова. Споделената икономика – принципи и технологии	148
20. Катя Рашева-Йорданова & Стефка Толева-Стоименова & Димитър Христозов & Веселин Чантов & Боряна Николова. Опит и експертност при специалистите в областта на научните данни	156
БЕЛЕЖКИ НА РЕЦЕНЗЕНТА	165

EXPERIENCE AND EXPERTISE OF DATA SCIENCE AREA SPECIALISTS

Senior Assist. Katia Rasheva-Yordanova, PhD¹ &

Senior Assist. Stefka Toleva-Stoimenova, PhD² &

Prof. D.Sc. Dimitar Christozov³ & Assoc. Prof. Veselin Chantov, PhD⁴

& Boriana Nikolova, PhD student⁵

University of Library Studies and Information Technologies

e-mail: k.rasheva@unibit.bg¹ & s.toleva@unibit.bg² & d.bristozov@unibit.bg³

& v.chantov@unibit.bg⁴ & b.nikolova@unibit.bg⁵

Abstract. *We live in the age of big data where dynamic new businesses are powered by different types of data. Virtually everyone using a smartphone is actually equipped with a data-generating sensor platform. In response, labor needs have quickly changed.*

In this situation, the in-depth analysis of the current changes in labor market due to the specifics of these data is almost a must. Today it is important to understand the available storage, processing and searching capabilities of large datasets, but more importantly there must be skills how to extract the useful knowledge from the data and how to use that knowledge. This necessitates the emergence of specialists with the necessary expertise, qualification and experience. This paper presents the set of the required skills that a modern data specialist must have in order to meet the expectations of the business. The focus of the study is to reveal the available levels of independence in the Data Science profession and to differentiate the training programs according to the needs and skills of the students.

Keywords: *Data Science, Employers, Junior Data Scientist*

Introduction

We live in the age of big data where dynamic new businesses are powered by different types of data. Virtually everyone using a smartphone is actually equipped with a data-generating sensor platform.

Today it is important to understand the available storage, processing and searching capabilities of large datasets, but more importantly there must be skills how to extract the useful knowledge from the data and how to use that knowledge.

This phenomenon has allowed the term Data Science to attract significant attention in the last few years turning specialists in Data Science into important experts. The role of such specialists has been widely discussed and academically acclaimed. An IBM analysis has shown that “by 2020 the number of positions for data and analytics talent in the United States will

increase by 364,000 openings, to 2,720,000. In 2020, job openings for data scientists and similar advanced analytical roles will reach 61,799. This is a significant number, but it represents just 2% of the projected demand across all job roles requiring data and analytics skill“ [14].

A constant growth is observed in the business need for qualified specialists with the necessary expertise to learn from big data. According to Gupta and George [8], the availability of skills for working with large data is a potential source of competitive advantage, which in turn is a critical resource for the accumulation of measurable business value [1]. The notion that companies need to combine technology with human skills traces its roots in the socio-technical framework. This enhances the role of human skills and states that maximum technological performance requires both management and nurturing of human skills and knowledge of organizations [10]. This notion is specially important in the context of big data, because skills are not only set on the exploitation of technical resources (such as software and infrastructure). More important is the generation of an insight that guides organizational decisions [12].

Studying big data can provide the answers to key challenges and help make more effective decisions based on evidence. Opportunity is then provided to convert complex and often unconstructed data into applicable information as a strategic response to the changing global trends [3].

Mikalef, et al. [9] claim that data literacy is probably the most sought by companies with big data, since the skills comprised in the scientist profile allow the companies to ask the right questions and turn data into practical visions. They come to the conclusion that software, infrastructure, and data are insufficient to provide any value unless personal skills and knowledge are available to implement them. Similar findings have been noted in a number of studies.

Data Scientist is a multidisciplinary profile that seeks knowledge in several areas of learning. According to Manieri [2] the Data Science Professional is an expert that has the ability to manipulate and retrieve knowledge turning it into a meaningful value. This specialist relies heavily on the scientific way of doing things, due to which his/her research experience is of high significance.

What is clear is that business is looking for those experienced scientists who have the ability not only to manipulate vast amounts of data with advanced statistical and visualization techniques but also to have a solid insight from which they can derive prospects. These insights help to predict potential outcomes in order to mitigate potential threats to business.

This paper presents the set of the required skills that a modern data specialist must have in order to meet the expectations of the business. The focus of the study is to reveal the available levels of independence in the Data Science profession and to differentiate the training programs accord-

ing to the needs and skills of the students.

The paper is organized in two relatively independent sections. The first section reviews the necessary knowledge, skills and competences that the modern data specialist should have where the emphasis is on the required ones. The second section presents ideas for differentiation of the Data Science training programs, according to the input skills and needs of the learners.

1. Compulsory Competences in the ERA of Big Data

Contemporary studies [11] indicate that data scientists are highly-educated – 88% of them are MAs and 46% are PhDs. With a few exceptions, usually a high level of education is necessary to develop the depth of knowledge needed to reach the scientist's skills. Most often, data scientists have a BA degree in computer science and statistics.

The distribution of the specialists according to their specialty shows that the largest number of specialists have graduated in Mathematics and Statistics (32%), followed by Computer Science (19%) and Engineering (16%). It is assumed that each of these specialties can provide the skills that a specialist who processes and analyzes big data needs.

Data Science is a dynamic area requiring constant updating of the acquired knowledge, skills and competences. Data scientists are expected to know a lot – machine learning, computer science, statistics, mathematics, data visualization, communication, and deep learning. Within those areas there are dozens of languages, frameworks, and technologies data scientists could learn [5].

In order to reflect the compulsory competencies that a modern data specialist should have, a review of the active job ads on the most popular online job sites in Bulgaria was carried out. Emphasis was placed on the knowledge, skills, experience and the requirements of proficiency in programming languages and tools sought by businesses.

The Job Listings Review showed that the business skills sought can be grouped into common categories, including analysis, machine learning, statistics, computer science, communication, mathematics, AI, etc. (Fig. 1)

The results show that the analysis and machine training are major skills building the data scientist's profile. Retrieving insights from data is the major function of data science and machine training includes skills for forecasting. The Age of Big Data has transformed the tools and skills employees utilize and workers across job functions now require analytical aptitude [14].

What is interesting is that communication skills are required in more than half of the job advertisements. The data specialists should be able to convey ideas and work in a team. Although not appearing so often in job advertisements, AI and deep learning can be considered as adjacent to machine learning. This is due to the fact that deep learning is increasingly used in machine learning and in natural language processing.

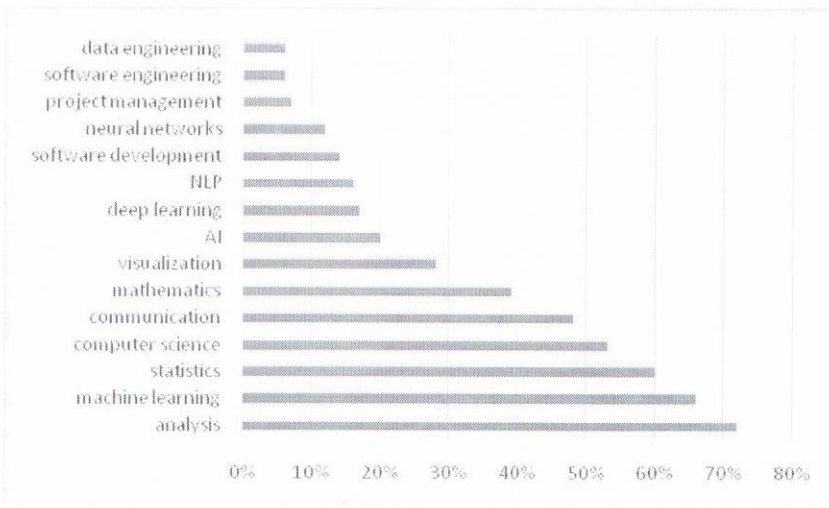


Fig. 1. General Skills in Data Scientist Profile

In most job advertisements, employers require skills to work with specific software tools as well as knowledge of working with different programming languages, libraries and tech tools.

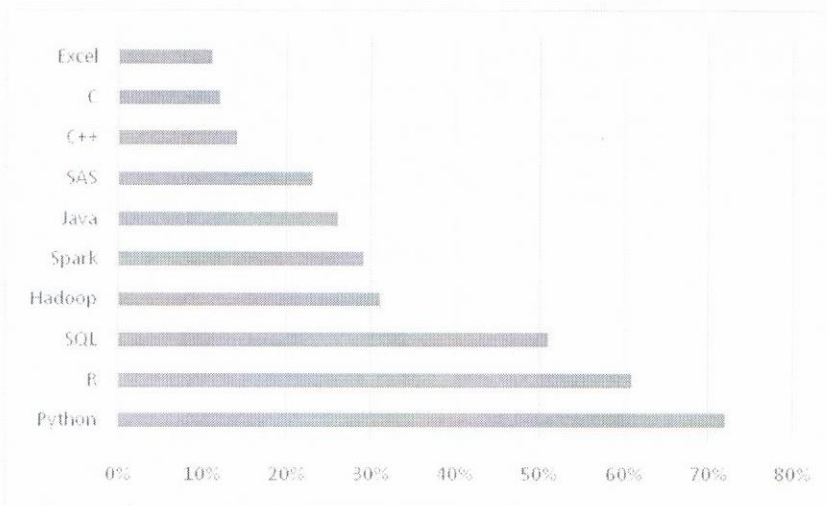


Fig. 2. Technology skills in Data Scientist Job

Python turns out to be the most sought programming language. The popularity of this open source language continues to grow. One reason is

that it is friendly to beginners and has a lot of support resources. Moreover, much of the data science tools are compatible with it. We can argue that today Python is the primary language that builds the data scientist's profile.

As can be seen in Fig. 2, the language R is not far behind Python. It was until recently thought that R is the basic data science language, but it has started to be gradually replaced by Python. The power of this open source language is in working with statistics, which is why it is still very popular with statisticians. However, both Python and R are a requirement for the position of the data scientist.

SQL is also listed in the Employer Specific Requirements List. SQL as a structured query language is the primary way to interact with relational databases. Recently, there seems to be a decrease in the demand for SQL by employers, yet the skills to work with this language are not negligible.

The above is confirmed by a study published online in *Towards Data Science*, according to which the representative of the data scientist is most commonly described as: "A male (the data science profession is heavily dominated by men, who accounts for 70% of the sample.), who speaks at least one foreign language and has a second-cycle academic degree (MA or PhD). He has been in the workforce for 4.5 years, after taking 2 years to land the title. R and Python are the preferred coding languages followed by SQL" [4].

On the basis of all this, we can summarize that the compulsory competencies for any modern data scientist include:

- **Analytical and critical thinking** – data scientists must be able to think critically in order to be able to apply an objective analysis of the facts about a topic or problem before formulating opinions or making judgments.
- **Coding** – First-rank data professionals know how to write code that is convenient in implementing different programming tasks.
- **Statistics and Mathematics** – The role of data scientists is to use their profound experience in mathematics to develop statistical models that can be used to develop or change key business strategies.
- **Communication skills** – has the ability to cooperate closely with business executives to communicate what is actually happening in the "black box" of complex equations in a way that reassures the business that it can trust the results and recommendations.

Based on a previous study [6] we can group these skills as follows: hardskills, softskills and analytical skills. The first group comprises coding, statistics and mathematics. Softskills comprise all communication skills and the analytical skills comprise the analytical and critical thinking.

Each of these skill categories shapes the profile of the data scientists. However, it is noticeable that the job advertisements are based on experience and expertise. An overview of this differentiation will be made in the next section of this article.

2. Experience and Expertise in Data Science Area

Data scientists use their analytical skills by operating freely with their hard skills. Based on the model hard-soft-analytical skills [6], the distribution of the tasks carried out in the sphere of each data science competencies can be presented as follows: (Table 1)

Table 1.

Structure of Data Science tasks [6]

Category	Tasks	Skills
Extract	Choose, Classify, Collect, Compare, Configure, Contrast, Define, Demonstrate, Describe, Execute, Explain, Find, Identify, Illustrate, Label, List, Match, Name, Omit, Operate, Outline, Recall, Rephrase, Show, Summarize, Tell, Translate	Hard skills
Verify	Apply, Analyze, Build, Construct, Develop, Examine, Experiment with, Identify, Infer, Inspect, Model, Motivate, Organize, Select, Simplify, Solve, Survey, Test for, Visualize.	Hard skills Analytical skills
Interpret	Adapt, Assess, Change, Combine, Compile, Compose, Conclude, Criticize, Create, Decide, Deduct, Defend, Design, Discuss, Determine, Disprove, Evaluate, Imagine, Improve, Influence, Invent, Judge, Justify, Optimize, Plan, Predict, Prioritize, Prove, Rate, Recommend, Solve.	Hard skills Soft skills Analytical skills

The Job Listings Review showed that companies need a data scientist with a variety of expertise and experience. Employers distinguish between Junior, Mid-Level and Senior Data Scientist and clearly distinguish their needs from the requirements specified to the candidates. The following table will compare the experience and education requirements for each of the three categories of specialists.

Table 2.

Category	Experience (years)	Educational level
Junior Data Scientist	From 0 to 2 years of experience in development / environment policy and statistics	Minimum of a Bachelor's degree in Statistics, Mathematics, or Computer Science
Mid-Level Data Scientist	3 or more years	Master's or a PhD degree
Senior Data Scientist	5+ years research experience	PhD in statistics or closely related field

The main difference between the three levels is related to the possibilities for self-reliance and independence in performing assigned tasks on the one hand. On the other hand, the difference is provoked by the possibilities of organizing the work process in accordance with the needs and the terms of the structure to which the respective specialists belong. More significant are the differences between Junior and Senior Data Scientists, therefore a comparison between these two levels will be made in this section.

Basically, Senior Data Scientists understand that ‘data’ always has flaws. These flaws can be data generating processes, biases in data [15]. Besides, Senior Data Scientists understand the “soft” side of making technical decisions. In practice, Senior Data Scientist is the specialist, on whom the organization and automation of tasks depends. This professional is responsible for outsourcing tasks to junior members or to consultants [13]. Part of the many other responsibilities of the Senior Data Scientists include [13]:

- Managing people, hiring the right people, managing managers who report to you.
- Training colleagues who might not be tech-savvy.
- Identifying the right tools and assessing the benefits and minuses of vendor software and platforms, for a specific large-scale project (construction of a huge taxonomy, etc.).
- Identifying the right algorithms and statistical techniques for a specific project.
- Identifying useful external or internal data sources, blending various data sources while cleaning data redundancies and other data issues.
- Understanding executive talk, and translating executive requests, questions, concerns, or ideas into successful data science implementations.
- Measuring the ROI that you bring to your company Interacting successfully with managers / colleagues / executives / clients of all kinds. Mostly a communication issue.

The competences listed are based on many skills and practices. The experience and expertise in each profession is developed in the process of its practice. When developing a training curriculum for an MA degree in Data Science, the profile of an expert with expertise is not attractive. For us, more interesting are the specialists in the Junior category. These are specialists who graduate from their undergraduate or MA degree program to become professionally trained – without experience and without experience.

As it became clear in the previous section, every data scientist is under pressure of the pace of development of new technologies, and in order to maintain the expertise, he or she must constantly build upon and develop their knowledge and skills. It is a good practice to differentiate qualification courses, training programs, etc. according to the level reached by the students or specialists. Based on the existing divide set by employers in job ad-

vertisements, it is necessary to target the training to each of the individual groups. Thus, students who have completed an MA program in Data Science will automatically fall into the Junior Data Scientist group because they do not have experience and experience. At the same time, practitioners need a different type of training – not so fundamental but purely practical.

An overview of Data Science training worldwide has shown that there are qualification courses aimed at an absolute beginner level and beginner to an advanced level. Such an upgrade in the university programs was not discovered.

The opportunities for this upgrade of skills to be implemented in universities will be the subject of further research.

Conclusion

The constant data generating and its accumulation in different storage databases requires the availability of specialists able to retrieve knowledge from them. An overview of job listings has shown that employers today distinguish between experts with expertise, qualifications and experience, and those who do not have them. Differences between Junior, Mid-Level and Senior Data Scientists in terms of employers' requirements call for measures to be taken to train data-handling professionals. We believe that building training courses and training programs tailored to the needs and input skills of the trainees is the foundation for developing quality professionals who meet today's business needs. Further research is to be done whether this will play such a role.

Acknowledgements

This work has been supported by National Science Fund at the Ministry of Education and Science, Republic of Bulgaria, within the Project DM 12/4 - 20/12/2017.

References

1. **Bharadwaj, A.** A resource-based perspective on information technology capability and firm performance: an empirical investigation. *MIS quarterly*, vol. 24, no. 1, pp. 169 – 196, 2000.
2. **Manieri, A., Demchenko, Y., Brewer, S., Hemmje, M., Riestra, R., & Frey, J.** (2015). Data Science Professional uncovered How the EDISON Project will contribute to a widely accepted profile for Data Scientists. In 2015 IEEE 7th International Conference on Cloud Computing Technology and Science Data. doi:10.1109/CloudCom.2015.57
3. **Daniel, B.** Big data and analytics in higher education: Opportunities and challenges, *Brit. J. Educ. Technol*, 46 (2015), 904–920.

4. **Valchanov, I.** What are the Skills Needed to Become a Data Scientist in 2018? // <https://towardsdatascience.com/what-are-the-skills-needed-to-become-a-data-scientist-in-2018-d037012f1db2>
5. **Hale, J.** The Most in Demand Skills for Data Scientists. What are Employers Looking For? Towards Data Science, 04.03.2019 // <https://towardsdatascience.com/the-most-in-demand-skills-for-data-scientists-4a4a8db896db>
6. **Rasheva-Yordanova, K., V. Chantov, I. Kostadinova, E. Iliev, P. Petrova, B. Nikolova.** Forming of Data Science Competence for Bridging the Digital Divide, 8th edition of the “The Future of Education” conference, PIXEL, Retrieved from https://conference.pixel-online.net/FOE/virtual_presentation_scheda.php?id_abs=3236
7. **Rasheva-Yordanova K., S. Toleva-Stoimenova, D. Christozov, I. Kostadinova.** Road Map in Developing Data Science Competences. Inted 2019 (in print).
8. **Gupta, M., J. F. George.** Toward the development of a big data analytics capability. *Information & Management*, vol. 53, no. 8, pp. 1049 – 1064, 2016.
9. **Mikalef, P. I. O. Pappas, J. Krogstie, and M. Giannakos.** Big Data Analytics Capabilities: a Systematic Literature Review and Research Agenda, *Information Systems and e-Business Management*, pp. 1 – 32, 2017.
10. **Wamba, S. F., A. Gunasekaran, S. Akter, S. J.-f. Ren, R. Dubey, and S. J. Childe.** Big Data Analytics and Firm Performance: Effects of Dynamic Capabilities. *Journal of Business Research*, vol. 70, pp. 356 – 365, 2017.
11. **Simplilearn,** 9 Must-have skills you need to become a Data Scientist, updated. Kdnuggets, 04.03.2019 // <https://www.kdnuggets.com/2018/05/simplilearn-9-must-have-skills-data-scientist.html>.
12. **Schoenherr, T., C. Speier-Pero.** Data science, predictive analytics, and big data in supply chain management: Current state and future potential. *Journal of Business Logistics*, vol. 36, no. 1, pp. 120 – 132, 2015.
13. **Granville, V.** 22 Differences Between Junior and Senior Data Scientists, *Data Science Central*, 04.03.2019 // <https://www.datasciencecentral.com/profiles/blogs/10-differences-between-junior-and-senior-data-scientist>
14. **Markow, W., S. Braganza, and Bledi Taska,** with Steven M. Miller and Debbie Hughes. *The Quant Crunch How the Demand for Data Science Skills is Disrupting the Job Market.* © Burning Glass Technologies 2017, p. 25, 2017.
15. **What Does it Mean to Be a Senior Data Scientist?, DATAQUEST,** 04.03.2019 // <https://www.dataquest.io/blog/what-does-it-mean-to-be-a-senior-data-scientist/>