World Journal on Educational Technology: Current Issues

Volume 12, Issue 1, (2020) 023-036

www.wj-et.eu

Learning skills for enhancing the use of Big Data

Jolanta Sabaitytė*, Faculty of Business Management, Department of Business Technologies and Entrepreneurship, Vilnius Gediminas Technical University, Sauletekio av 11, Vilnius LT 10223, Lithuania
Vida Davidavicienė, Faculty of Business Management, Department of Business Technologies and Entrepreneurship, Vilnius Gediminas Technical University, Sauletekio av 11, Vilnius LT 10223, Lithuania
Roberta Karpoviciutė, Faculty of Business Management, Department of Business Technologies and Entrepreneurship, Vilnius Gediminas Technical University, Sauletekio av 11, Vilnius LT 10223, Lithuania

Suggested Citation:
Sabaitytė, J., Davidavicienė, V. & Karpoviciutė, R. (2020). Learning skills for enhancing the use of Big Data. *World Journal on Educational Technology: Current Issues*. *12*(1), 023–036 https://doi.org/10.18844/wjet.v12i1.4438

Received August 15, 2019; revised from October 20, 2019; accepted from January 1, 2020.
Selection and peer review under responsibility of Prof. Dr. Servet Bayram, Yeditepe University, Turkey.

Abstract

For a competitive organisation, it is important to invest in employee's education to keep their knowledge up to date and ensure continuous growth in terms of employees' competence and skills. Continuous learning is considered as one of the success factors for the organisation, since this ensures constant growth of employees' competence that results in productiveness and effectiveness in its product development. Nowadays, more and more organisations are investing in performance analytics, Big Data analytics, artificial intelligence and process automation that are strongly related. However, such initiatives and projects can fail due to the lack of skills and competence of employees, which is strongly related to certain learning barriers. Therefore, the objective of the article is to create a conceptual framework for the elimination of learning-related barriers. The methods employed in this research are scientific literature analysis, synthesis and logical analysis of information, comparison of information, systemisation and visualisation. The conceptual framework illustrates the learning process for learning-related barriers elimination and suggests learning solutions: e-learning, mentoring, in house sessions, generic courses and external training solutions such as conferences and seminars.

Keywords: Teaching concept, learning big data, big data management, big data teaching, electronic learning.

* ADDRESS FOR CORRESPONDENCE: **Jolanta Sabaitytė,** Faculty of Business Management, Department of Business Technologies and Entrepreneurship, Vilnius Gediminas Technical University, Sauletekio av 11, Vilnius LT 10223, Lithuania. *E-mail address*: jolanta.sabaityte@gmail.com

1. Introduction

With the formation of the knowledge society, knowledge, information, data and its purposeful use and processing have become the basis for further development of society. This has led to the need to collect and process large amounts of data, which are now widely called – big data. The term 'big data' stands for modern technologies for processing data that is generated at high speed, in large volumes and a variety of structures (Bagriyanik & Karahoca, 2016; Lee, 2017; Turkben, Turkben, Karahoca, & Karahoca, 2016). Optimally processed Big Data creates a supportive information environment for organisations that enables them to get to know the customers. As a result, Big Data research does not only focus on academia but also on business (Beyer & Laney, 2012; Chanunan, 2017; Gartner, 2017; Langkafel, 2015; Marr, 2015; Pabedinskaitė, Davidavicienė, & Milicauskas, 2014). Sometimes companies outline artificial intelligence or process automation. However, it can be stated that all these three areas of performance strongly related and require similar competences.

For the successful functioning of the organisation, investment in human resources is important since employees need to keep their skills and knowledge up to date (Chang, 2016; Gezgin, 2018; Keser & Semerci, 2019). By offering employees training and working with learning processes, an organisation can increase employees' productivity and effectiveness, thus ensuring competence improvement continuously. This is a great advantage in such a competitive market since a well-organised learning process can result in growth for all organisations due to increasing employees' performance and better products or services development. For an organisation that is working with technologies such as Big Data, it is highly important to ensure continuous learning of its employees to extract the greatest value from collected data and successfully execute its projects (Smaliukiene & Giedraityte, 2018; Soubra, Tanriover, & Orhun, 2017; Tvrdikova, 2016). However, when planning learning processes, it is important to choose the right learning methods to achieve cost-effective and profitable results.

The problematic issue of this work is the learning concept, including learning techniques and related skills that are required for successful Big Data projects execution. Before starting Big Data implementation organisation should evaluate its resources and identify skills that are needed for successful Big Data project execution. The lack of knowledge and skills may have a negative impact on extracted data results with incorrect or unprofessional predictions and insights. Moreover, a poorly chosen learning strategy, as well as a lack of knowledge, can lead to employee resistance and inefficient use of time and cost. Such human-related barriers need to be eliminated before the started project execution. However, it is still unclear, which are the key skills that should be associated with professional data analytics and management. Therefore, the objective of the article is to create a theoretical framework for gaining competences and proposing learning programs necessary for successful Big Data project execution.

Literature and other references analysis, synthesis and logical analysis of information, comparison of information, systemisation and visualisation methods employed in this research.

2. The concept of Big Data and its evolution paradigm

The rising use of social media and advances in the Information and Communication Technologies sector results in a growth of the amount of data generated by individuals and organisations. This reflects in the daily creation of structured and unstructured data that is difficult to process and analyse with the existing information technology (IT) infrastructure and tools. Such data are called Big Data. Although the term Big Data is currently and widely used, there is no uniform definition in the scientific literature (Ward & Barker, 2013). Different authors formulate this concept from different angles, and there are different kinds of aspects they emphasise, for example:

- Alharthi, Krotov, and Bowman (2017) define large amounts of large-scale data created by individuals or organisations activity, which is difficult to process using traditional analytical tools.
- Politaitė and Sabaitytė (2018) claim that Big Data is large data arrays for which conventional data processing tools are not proper.
- As defined by Yin and Kaynak (2015), Big Data is datasets that exceed the capacity of regular database software to store, process and analyse.
- Sharma (2015) also emphasises that these are tremendous and complicated datasets that are difficult to process using regular data processing tools.

Some authors propose to treat Big Data as a massive amount of human-generated digital data that is complicated to process using traditional data analytics tools (Alharthi et al., 2017). Other authors use the 3V model to define the concept of big data, which defined the concept of Big Data in the early stages of Big Data development. The 3V model has three main features of big data: Volume, Velocity and Variety (Laney, 2001):

- *Volume*. This characteristic points out the amount of data that an individual or organisation collects or generates (Lee, 2017). Big Data is data in a variety of formats, no longer in gigabytes, but significantly larger units – petabytes, zettabytes or exabytes. As data warehouse capacities grow steadily and allow for the accumulation of ever-larger datasets, definitions of data volume vary according to factors such as time and data type (Gandomi & Haider, 2015). This means that what is seen as a Big Data today may no longer be rated as Big Data later.
- *Velocity* refers to the speed of data generation and processing (Sharma, 2015). Initially, organisations used batch processing systems to analyse data (when a document is opened, a program is called to process it), which made the data processing process slow and costly (Lee, 2017). However, data speeds and the use of digital devices have accelerated this process over time. This has made it easy for various businesses to perform Big Data analytics in real time.
- *Variety* refers to the number of data types. Data are collected from a diversity of sources like spreadsheets, databases, text files or digital data streams (Gartner, 2017). As a result, Big Data is seen as arrays of data in different formats. These arrays are made up of structured, semi-structured and unstructured data. Structured data are tabular data that can be found in spreadsheets and databases, when text, photos, audio and video files stand for unstructured data; Extensible Markup Language is an example of semi-structured data (Gandomi & Haider, 2015).

In newer literature, the authors broaden the concept of Big Data with 5V model, which is complemented by a 3V model with two new components – Veracity and Value (Sharma, 2015; Yin & Kaynak, 2015):

- *Veracity* is important as most data coming from the internet or various sensors are incorrect (Rajaraman, 2016). Data collection should include requirements that define data accuracy and correctness. Therefore, it is important that erroneous data would be removed before the analysis processes begin. The company should define what data are needed and for what purpose, thus reducing the amount of false or unnecessary data (Yin & Kaynak, 2015). Data flows are growing rapidly, and large volumes are more difficult to process, requiring more resources and leading to misconceptions.
- Analysing the collected data provides valuable insights on how companies can adjust their strategy, improve their production and gain a competitive edge. This leads to another characteristic, which is *Value*. It is important to understand that the data itself are not valuable; the value is obtained by analysing it (Rajaraman, 2016). Cavanillas, Curry, and Wahlster (2016) provided the data value creation chain, which contains the processes of data collecting, analysis, mentoring, storage and use (see Figure 1). As a data analyst can track business trends in a variety of queries and filter the data, the value can be described as a key feature of Big Data (Hashem et al., 2015; Sivarajah, Kamal, Irani & Weerakkody, 2017) and the final product.

Figure 1. The chain of Big Data value creation (created by authors based on Cavanillas et al., 2016)

Analysing the concept of big data, 6V (Groves, Kayyali, & Knott, 2013) and 7V (Gandomi & Haider, 2015) models with Validity and Visibility characteristics were offered:

- *Validity* – this characteristic describes the logic of the data (Khan et al., 2014). At first sight, this characteristic may be associated with veracity. However, when speaking about validity, it is more important whether the data have the meaning assigned to it or whether it corresponds to the facts. The need for justification can be linked to the protection of personal data, as it is necessary to ensure that the source of the data is transparent and appropriate for storing such data.
- *Visibility* is also important to establish that the data are authentic and accessible. This feature allows the user to see and analyse the data with their access (Gandomi & Haider, 2015).

V models are evolving – they are complemented by new elements based on Big Data properties. Considering that the V models are built on the characteristics of big data, they can be treated as models of Big Data characteristics (see Figure 2).

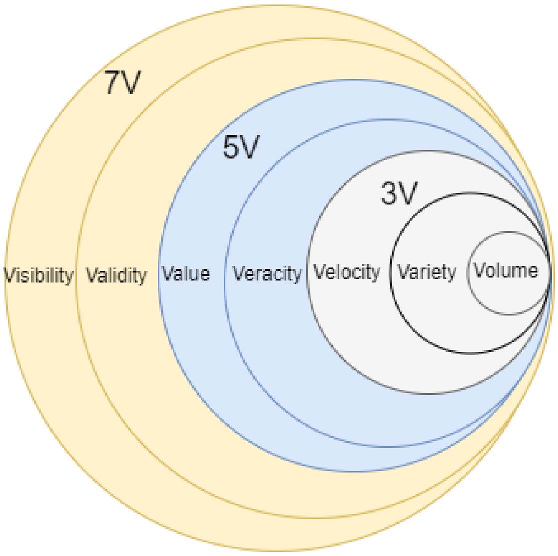

Figure 2. Interaction between the Big Data models (compiled by authors based on Alharthi et al., 2017; Gandomi & Haider, 2015; Khan et al., 2014; Rai et al., 2018; Russom, 2011; Sharma, 2015; Tanwar, Duggal, & Khatri, 2015; Yin & Kaynak, 2015)

Furthermore, the evolution of these models and their characteristics has shown that the concept of Big Data is evolving rapidly. However, such sudden changes in cognition have caused a great deal of confusion. This is confirmed by a study of 154 executives commissioned by SAP at Harris Interactive (2012). During the survey, corporate executives were asked to define the concept of big data. About 28% identified Big Data as a strong growth in transactional data; 24% as a new technology to address the high volume, speed and diversity of data; 19% stated that Big Data defines the requirements for storing and archiving data and 18% treated Big Data as new data origins referring to social media and mobile devices (SAP, 2012). The study has shown how different companies' managers interpret the concept of Big Data differently. Some executives focus on what Big Data is, whereas others focus on what that data does. The problem is that managers are not familiar with the concept of big data.

Moreover, the concept of Big Data is also interpreted in a broad sense as a phenomenon that is affected by various factors. Boyd and Crawford (2012) describe Big Data as a cultural, technological and scientific phenomenon based on:

- *Technology*: Computing, analysing, linking and comparing large datasets maximise computing power and algorithm-based accuracy.
- *Analysis*: Big Datasets are used to determine economic, social, technical and legal requirements.
- *Myths*: There is a belief that large datasets are encoded with information that provides objective and accurate insights that cannot be obtained from other sources.

Other authors focus on the field of business intelligence, which is a consequence of Big Data. Chen, Peng, and Lee (2012), while defining big data, also identify related technologies: cloud computing, Internet-enabled devices and Hadoop Big Data processing software. As business intelligence is expanding rapidly and increasing amount of companies is collecting and analysing big data, security and privacy concepts have become closely associated with big data. Therefore, it is appropriate to further analyse the risks of using big data.

To sum up, it can be said that Big Data is a large variety, high-volume and fast-growing data that require advanced storage and processing technologies and that can provide valuable insights. However, it is important that data would meet conditions of fairness, reasonableness and visibility since these are key features that can ensure data quality and value extraction. After the identification of the main Big Data characteristics, it is crucial to analyse the risks of Big Data usage.

3. Barriers of Big Data technologies implementation

For a better understanding of the possibilities when using big data, it is also important to define the potential risks of this phenomenon, knowledge of which will provide the basis for more efficient application. Although collecting Big Data is becoming easier, 60% of such data projects still fail (Gartner, 2015). This is a result of the emergence of Big Data risks.

Some of the authors (Lee, Kao, & Yang, 2014; Wang & Li, 2015) mention that Big Data is inseparable with Industry 4.0 technologies, which is expected to radically change how businesses compete and operate. Current developments include Big Data inside and outside the company: structured and unstructured, computer, web and mobile data that enable object analysis to provide a historical overview and future projections. However, the success of the implementation of Big Data analysis into the organisation processes depends on the elimination of potential risks. Therefore, it is significant to identify and analyse the risks of using big data.

In the initial phase of Big Data analysis, there is a risk that the results will not be accurate. This is an essential risk when considering the risks of using Big Data in the broadest sense. This risk manifests itself through the 5V model, which has five characteristics that describe big data: *volume, variety, velocity, veracity* and *value*; there are many cases where known analytical tools cannot be applied to large amounts of data in various formats (diversity) within a reasonable time (low speed), resulting in inaccurate results (correctness) leading to false predictions (low value) (Krasnow Waterman & Bruening, 2014). For more accurate results and more accurate predictions, special systems and algorithms should be used to process Big Data (ur Rehman et al., 2016), which would guarantee faster, more reliable processing of complex data.

Analysing Big Data risks in the narrow sense, the underlying technological, human and organisational barriers are identified (see Figure 3).

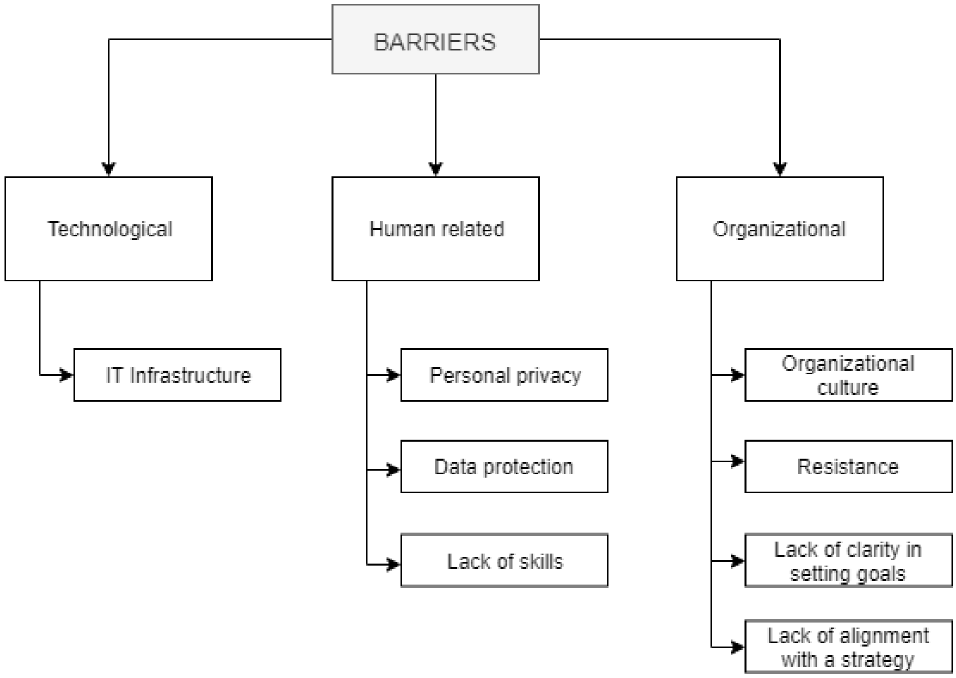

Figure 3. Classification of Big Data barriers (compiled by authors based on Alharthi et al., 2017)

Technological barriers are generally defined as IT infrastructure. Most IT applications currently in use cannot withstand the growing demands of Big Data analytics (Alharthi et al., 2017), which are changing rapidly. As a result, there is a risk that in-house technologies will not be adapted for Big Data analysis. Application of Commodity Hardware (Alharthi et al., 2017) helps to reduce this risk. For example, to store ever-increasing amounts of data, it is possible to expand the total capability of a database by combining several hard disks. As the company develops its existing IT infrastructure, there is another risk – poor allocation of financial resources. It is useful to plan investments for the interim period so that new uses of Big Data can be anticipated which may require new ways and means of data processing.

Human-related barriers are associated with a person's activities that affect the use of big data. Recently, great attention was paid to data protection and personal privacy. For example, location providers may identify the user by following their GPS information, which may be related to their home or work location information (Sivarajah et al., 2017). There is a risk of invading the privacy of the individual. To eliminate these risks, legal regulation should be the base for implementing big data-based technologies in organisations. Although analysing Big Data protection, malware is identified as a growing risk for businesses and treat for hackers (Abawajy, Kelarev, & Chowdhury, 2014). The use of such equipment enables data to be leaked, modified or deleted. Therefore, security controls that prevent malicious software are needed to protect stored data. In addition, it is very important to focus on the skills and competencies of the staff involved in Big Data Analysis. Most companies face a shortage of employees with Big Data and analytical skills (Tole, 2013). This puts data security at risk. Furthermore, the data analysis is of poor quality, and the results are erroneous.

The organisational culture of the company is illustrated by the strategy developed by the management and followed by all employees. It can be closely related to organisation barriers. There is a significant risk that organisational culture will interfere with the use of Big Data in the enterprise. Therefore, the required cultural changes must be organised. In order to achieve a successful cultural change, it is required to have a clear strategy for big data analysis. Employees need to know the benefits of big data, and what to expect from the results.

When discussing Big Data challenges, it can be noticed that it is closely related to general organisation resources. Gupta and George (2016) propose seven resources that allow organisations to create Big Data capability. Tangible resources stand for data, technology and other basic resources, including time, investments, human resources and managerial and technical skills. Data-driven culture and organisational learning are seen as two of the most important resources that are required for building Big Data analytics capability in the organisation. Waller and Fawcett (2013) also describe analytical skills and business and management understanding as to the most important skills for academic and applied professionals. Therefore, the development of Big Data specialists requires close cooperation between educational institutions and businesses, ensuring that theoretical knowledge meets practical needs (Miller, 2014). The organisation must first encourage decision-making based on the analysis of such data as this would help to avoid barriers related to the use of big data. Furthermore, implement appropriate Big Data processing technologies and train employees to understand and analyse big data, ensuring implementation of technologies alignment with the IT strategy of the organisation and general strategy. As well as choose the appropriate Big Data analysis methods and models to process the data collected by the organisation and decide whether the company needs to analyse the Big Data at all. This leads to the importance of learning and education for Big Data projects understanding and successful development.

4. Skills required for successful Big Data projects execution

When identifying potential risks and barriers related to Big Data projects, it can be noticed that some of the barriers such as organisational culture or lack of employers' skills are coming from the main body of the organisation. Such barriers are closely related to each other. For instance, lack of clarity in setting goal or resistance can be a result of lack of skills or knowledge which are required for understanding how Big Data can be used and what value it can bring for the organisation.

Guttmann (2018) research which was published on Statista data shows that the leading in demand skills related to data marketing as of 2017 was data analytics with 87.30% of respondents considering it as the most important competency in support of their future data-driven marketing projects, data management and processing was assumed as second most important competency considered by 81% of respondents (Figure 4). This shows that there is a huge demand for competent specialists who would be able to analyse Big Data and manage Big Data-related processes.

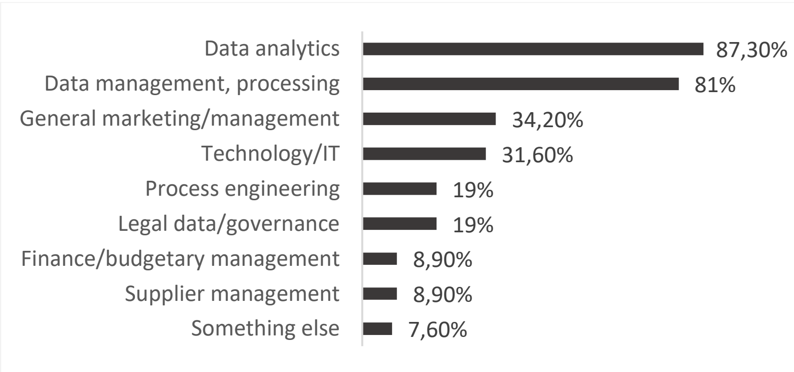

Figure 4. Leading in demand skills related to data marketing in North America as of November 2017 (Guttmann, 2018)

According to Gandomi and Haider (2015), Big Data analytics can be described as the processing and analysing Big Data in order to discover information within the Big Data. However, as it was previously stated, the value of Big Data is obtained by analysing it. Thus, employees that are working with such projects but do not have enough competence may fail in finding real value in analysed data and delivering insights. In such cases, tangible resources such as technology, time and investment would be wasted without delivering expected results.

Homer (2001) emphasises the importance of people skills for a company, as they impact every aspect of the organisation process. Thus, it is important and worth investing in required skills development and elimination of knowledge gaps. Although there is no specific research that could identify all competencies required for data analytics, there are several articles that are giving an abstract designation of competences. Gupta and George (2016) identify two groups of skills which are required for data analytics:

- *Technical skills* in Big Data projects stand for understanding the use of technologies and the ability to extract valuable insights from data. It is related to the ability to work with such Big Data processes as data extraction and data cleaning as well as competencies in machine learning and understanding such programming paradigms as MapReduce (Russom, 2011).
- *Managerial skills* are related to both the organisation's management and understanding where and how to apply the insights extracted from Big Data as well as being able to understand and predict both current and future organisation's needs (Gupta & George, 2016).

Furthermore, when introducing skill sets for data scientists, Schoenherr and Speier-Pero (2015) provided data management and enterprise business processes and decision making as two separate skill sets when Gupta and George (2016) introduced them as Managerial skills in general by showing that they are closely related. Skills needed to analyse Big Data based on specific features and abilities that can be described as (Rao, 2014; Dubey & Gunasekaran, 2015):

- *Hard skills* are domain and technical skills, e.g., statistics, forecasting, optimisation, quantitative finance, financial accounting, and so on.
- *Soft skills* are also known as interpersonal skills, e.g., leadership ability, team skills, listening skills, learning skills, positive attitude, communication skills, interpersonal skills, patience and passion emotional intelligence.

Debortoli, Muller, and vom Brocke (2014) provided a more detailed approach by identifying Big Data required competencies and organising them. Based on the authors' literature research, the below Big Data competencies were identified:

- *Domain skills* – sales, life science, digital marketing competencies;
- *Management* – startup, sales and business development competencies;
- *Concepts* and *Methods skills* – quantitative analysis, machine learning, database administration, software engineering, software testing, data warehousing;
- *Product skills* – NoSQL databases, Java, .NET and PHP/JavaScript programming languages. Furthermore, when introducing programming languages that are in demand for Big Data analytics, it is necessary to mention Python programming language, which according to Pedregosa et al. (2011) is a language of choice for machine learning.

It is important to notice that the above-mentioned competencies are related to hard skills. However, both hard skills and soft skills are important for Big Data projects execution, because a highly skilled professional with technical competence might not be able to communicate his insights or value of data for other employees without such soft skills as team skills or communication skills. It is important to understand that for successful project execution there should be both sets of skills – hard and soft skills combined.

Based on the researches and analysis, a theoretical model illustrating a set of skills required for successful Big Data project execution was created (Figure 5). The model illustrates hard and soft skills in demand for data analysis and data analytics related to project management.

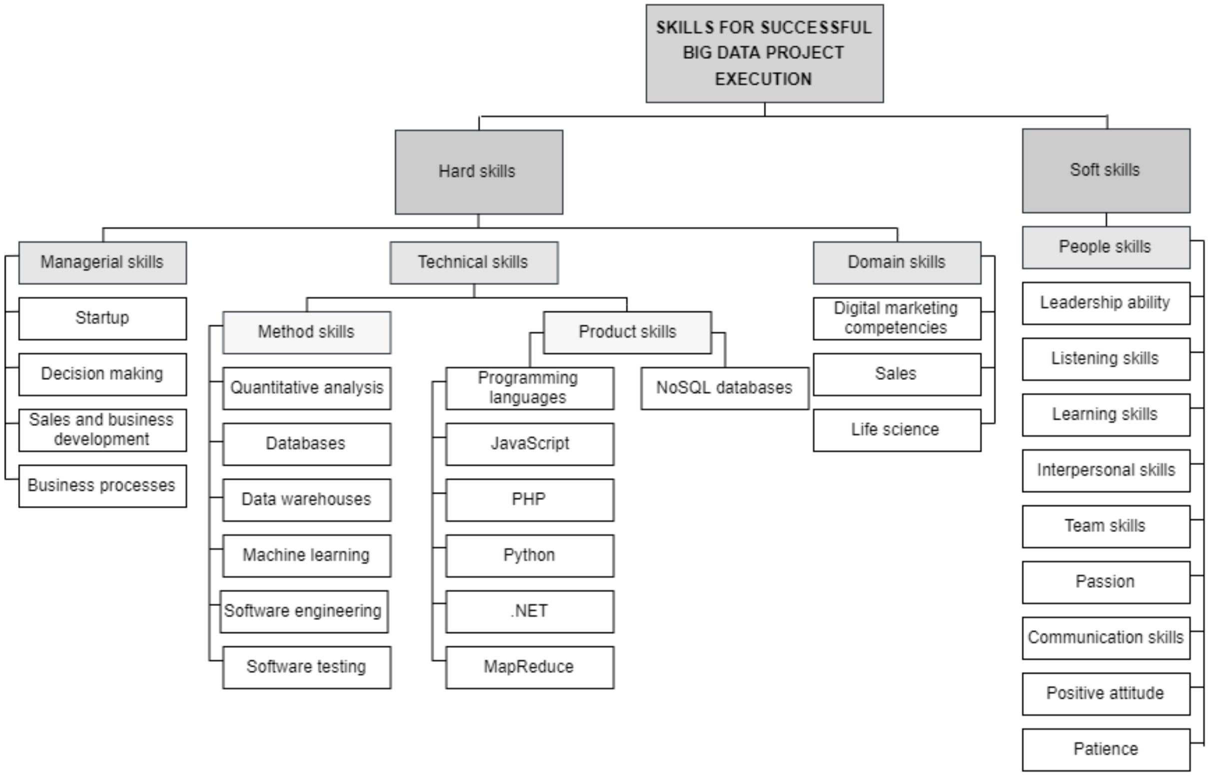

Figure 5. Skills required for successful Big Data projects execution (compiled by authors based on Debortoli et al., 2014; Dubey & Gunasekaran, 2015; Gupta & George, 2016; Pedregosa et al., 2011; Russom, 2011; Schoenherr & Speier-Pero, 2015)

Dubey and Gunasekaran (2015) provide a theoretical framework for Big Data skills education and training, which shows that both hard and soft skills can be learned by both formal and informal education and training. This gives an approach that both hard and soft skills can be developed through learning processes, and the barriers related to these skills can be considered as learning-related barriers that can be eliminated through the learning and education process.

5. Learning process analysis

Before starting employees', education and learning processes it is most important to identify skills that are required in order to achieve organisation goals, as this enables the organisation to proceed with more efficient training (Homer, 2001). Skills gaps can be identified by applying analysis of skills required for job or project execution and comparing it with employees' skills. For employee skills measurement, Robinson, Sparrow, Clegg, and Birdi (2007) proposed a three-phase methodology including preliminary interviews, questionnaires and critical incident technique interviews that could be used. Furthermore, skills management tools such as skills tracking solutions (e.g., TrackStar) can be used. After skills gaps are identified, employees' profiles should be revised for assessing their competencies. This would help to organise more cost-effective training by focusing on employees' groups which have gaps in specific skills. After skills gaps for employees are identified, they should be matched to training solutions. For individual learning such as learning options as e-learning or instructor-led, mentoring learning could be useful solutions. Navimpour et al. (2015) state that the web is a learning tool that allows us to build and share knowledge. Thus, nowadays, e-learning is appearing as a highly used learning attitude utilised by many organisations. This type of learning is cost-effective, and it is a great solution for continuous education (Chen, 2014). Furthermore, this type of learning is flexible, since employees can control their learning process with training various skills based on their competence and needs. In this type of learning, materials required for learning are available on an internationally accepted platform. For example, such platforms as Udemy that has a wide range of online courses that could be used for learning such skills as machine learning understanding, use of data analytics tools such as Hadoop and programming languages. Another option for employees' education can be mentoring-based learning. According to Klinge (2015), traditional mentoring can be associated with a senior person assisting a younger person's career development by counselling and sharing the feedback. Mentoring-based learning can be considered as a part of knowledge sharing organisational learning encouraging employees' cooperation and collaboration. Moreover, nowadays, platforms like SharpestMinds can offer remote mentorships that could be matched with challenge-based learning while focusing on current real-world problems as well as making the learning process more adjustable based on employees' needs. It can be noticed that nowadays technologies are giving great value for the learning process by extending the boundaries of learning possibilities. For team learning, there is a possibility to organise generic courses run by training companies as well as do in-house sessions. External training options such as seminars or conferences should be considered as a part of the training plan as well.

According to Gupta and George (2016) asserts that the ability to reconfigure organisations' resources can be resulted by its involvement in organisational learning. Therefore, in such cases, learning and organisational learning might be potential factors for problem-solving and internal growth in terms of employers' competence and skills. When learning is related to individual development processes, the benefits of organisational learning can be extracted through different angles which are used in definitions of organisational learning concept:

- Sattayaraksa and Boon-itt (2015) describe organisational learning as organisational processes that acquire, share, develop, utilise and store knowledge in order to achieve better organisational performance.
- The collaborative learning process of individuals (Song, Joo, & Chermack, 2009).
- The development of new knowledge that can potentially be used to influence an organisation's behaviour (Serinka et al., 2014).
- A learning process through social interactions at the groups and organisation levels (Bratianu, 2015).

It is crucial to develop a learning-oriented organisational culture, where learning and training would be encouraged, and employees were able to share knowledge and experiences. The solutions used for such knowledge sharing could be such activities as discussions, presentations or workshops that could lead to organisational learning.

Overall, organisational learning stands for the processes of learning and knowledge sharing in order to ensure better organisation performance and behaviour. That's why organisational learning is considered a competitive advantage and should be involved as a part of an organisation's culture. As organisational learning is focused on learning from experience and knowledge sharing, while being a part of organisation culture it can be an advantageous feature leading to the growth of the organisation and the successful project implementation and execution process. Based on the learning process analysis, a conceptual framework for learning-related barriers elimination was created (Figure 6). The framework illustrates a process that is leading the organisation to continuous learning to ensure learning-related skills barriers elimination for successful big data related to project execution.

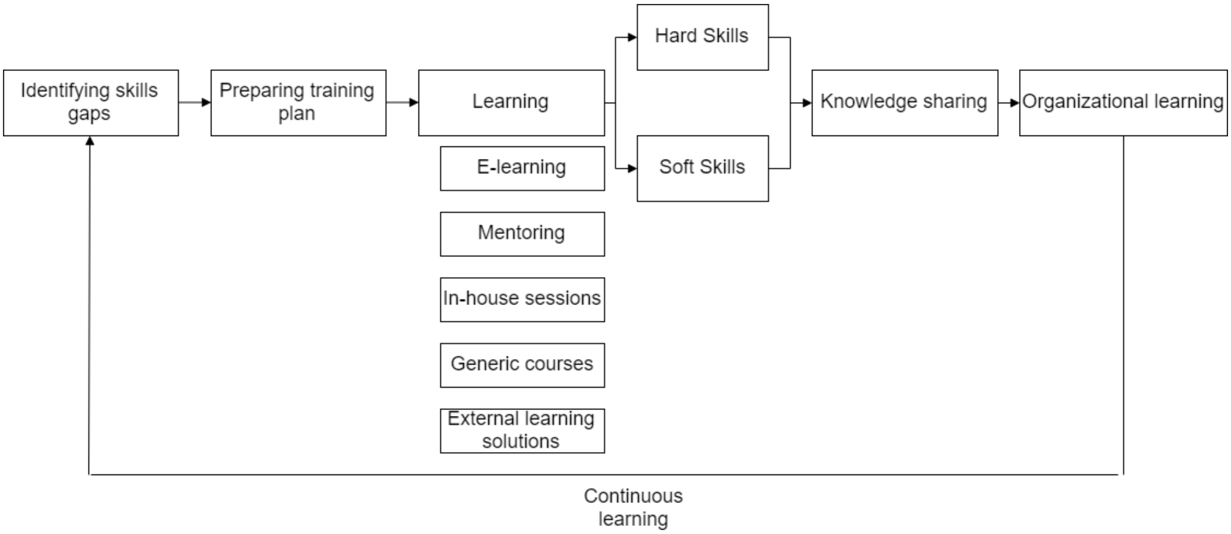

Figure 6. The conceptual framework for the learning process (compiled by authors based on Chen, 2014; Gupta & George, 2016; Jafari Navimipour & Zareie, 2015; Klinge, 2015)

6. Conclusions

The analysis of scientific and other literature allowed to identify the concept of Big Data has shown that there is no single definition to describe big data. However, most authors define Big Data according to their attributes such as volume, velocity, variety and so on. Summing up the Big Data can be defined as the volume of high-volume and high-speed information generated by people or devices and requiring innovative and evolving technologies capable of collecting, storing and processing them. Big Data is diverse and can be categorised based on different features, such as source, content, collection, storage and processing. The most important categories of data in the literature are structured, unstructured and semi-structured data, based on data content and format. It was found to be a complex concept that is still evolving. It should be emphasised that the authors define the concept of Big Data in a narrow and broad sense and emphasise the different characteristics of such data. This expands the concept of Big Data and causes a lot of confusion in the cognitive sense.

Nowadays, organisations that are analysing Big Data have a huge competitive advantage since they have the ability to understand their customers' behaviour better. With the use of Big Data insights, an organisation can plan its strategy more efficient and achieve better results. However, since the main value can be extracted from data only through advanced analytics it is important that employees that are working with Big Data would be highly skilled to able to extract value from collected data.

The model shows the skills which are in demand for Big Data analytics. These skills contain both hard skills that are related to the domain, technical and managerial skills and soft skills that are considered as people skills. Since both hard and soft skills can be learned through training processes, these skills can be considered as learning-related skills. The lack of such skills is causing learning-related barriers that can be eliminated through organising both individual learning and organisational learning. Organisational learning is highly recommended as a factor that should be a part of an organisation’s culture since it is one of the factors that are driving the organisation to success.

The conceptual framework illustrates the learning process for learning-related barriers elimination. The model shows learning solutions: e-learning, mentoring, in house sessions, generic courses and external training solutions such as conferences and seminars, that can be considered for the training plan. While applying the above-mentioned learning solutions, both hard and soft skills, which required for successful Big Data project execution, can be learned. By knowledge sharing based on trained skills organisations can develop organisational learning culture leading to continuous learning resulting in learning-related barriers elimination and efficient Big Data implementation

7. Recommendations

When working on Big Data projects, it is important to have a clear vision of what the organisation wants to achieve by processing and analysing data. This requires an understanding of Big Data itself and the value that can be brought by it. Before starting such a project, the communication barriers related to benefits that can be delivered by data extraction should be eliminated. This would help all employees to be on the same page in strategy, and furthermore, it can work as a motivator for self-learning and education. However, one of the main challenges appearing when developing Big Data projects is learning-related barriers that require employees' education. These barriers can be eliminated by organising employees' learning processes and working on the intensity of cultural learning. When organising employees' education, it is recommended to follow the process as it can help to achieve more efficient results. Identifying skills required to achieve organisational goals, revising employees' competence and matching them with training solutions – these steps can help to organise a more effective learning process. Therefore, the well-organised learning process will be less time consuming as well as cost-effective. As was previously stated, organisational success is closely related to organisational learning intense, therefore it is highly recommended to involve learning processes in an organisation's culture.

References

Abawajy, J. H., Kelarev, A., & Chowdhury, M. (2014). Large iterative multitier ensemble classifiers for the security of big data. *IEEE Transactions on Emerging Topics in Computing*, *2*(3), 352–363.

Alharthi, A., Krotov, V., & Bowman, M. (2017). Addressing barriers to big data. *Business Horizons*, *60*(3), 285–292.

Bagriyanik, S., & Karahoca, A. (2016). Personal learning environments: a Big Data perspective. *Global Journal of Computer Sciences: Theory and Research*, *6*(2), 36–46. https://doi.org/10.18844/gjcs.v6i2.1474

Beyer, M. A. & Laney, D. (2012). *The importance of 'big data': a definition*. Stamford, CT: Gartner.

Boyd, D., & Crawford, K. (2012). Critical questions for big data: provocations for a cultural, technological, and scholarly phenomenon. *Information, Communication & Society*, *15*(5), 662–679.

Bratianu, C. (Ed.). (2015). *Organizational knowledge dynamics: managing knowledge creation, acquisition, sharing, and transformation*. Hershey, PA: IGI Global.

Cavanillas, J. M., Curry, E., & Wahlster, W. (2016). *New horizons for a data-driven economy: a roadmap for usage and exploitation of big data in Europe*. New York City: Springer International Publishing.

Chang, V. (2016). Review and discussion: e-learning for academia and industry. *International Journal of Information Management*, *36*(3), 476–485.

Chanunan, S. (2017). Emerging educational technologies in higher educational institutions: the current trends and impacts from Thailand universities' perspectives. *New Trends and Issues Proceedings on Humanities and Social Sciences*, *4*(1), 89–103. https://doi.org/10.18844/prosoc.v4i1.2239

Chen, T.-L. (2014). Exploring e-learning effectiveness perceptions of local government staff based on the diffusion of innovations model. *Administration & Society*, *46*(4), 450–466. https://doi.org/10.1177/0095399713482313

Chen, Y.-C., Peng, W.-C., & Lee, S.-Y. (2012). Efficient algorithms for influence maximization in social networks. *Knowledge and Information Systems, 33*, 577–601. https://doi.org/10.1007/s10115-012-0540-7

Debortoli, S., Muller, O., & vom Brocke, J. (2014). Comparing business intelligence and big data skills. *Business & Information Systems Engineering*, *6*(5), 289–300.

Dubey, R. & Gunasekaran, A. (2015). Education and training for successful career in Big Data and Business Analytics. *Industrial and Commercial Training*, *47*(4), 174–181. https://doi.org/10.1108/ICT-08-2014-0059

Gandomi, A., & Haider, M. (2015). Beyond the hype: Big Data concepts, methods, and analytics. *International Journal of Information Management, 35*(2), 137–144. https://doi.org/10.1016/j.ijinfomgt.2014.10.007

Gartner. (2015). *Gartner says business intelligence and analytics leaders must focus on mindsets and culture to kick start advanced analytics*. Retrieved from https://www.gartner.com/newsroom/id/3130017

Gartner. (2017). *Big data*. Retrieved from http://www.gartner.com/it-glossary/big-data/

Gezgin, D. M. (2018). Understanding patterns for smartphone addiction: age, sleep duration, social network use and fear of missing out. *Cypriot Journal of Educational Sciences*, *13*(2), 166–177. https://doi.org/10.18844/cjes.v13i2.2938

Groves, P., Kayyali, B., & Knott, D. (2013). *The 'big data' revolution in healthcare*. Retrieved from https://www.mckinsey.com/industries/healthcare-systems-and-services/our-insights/the-big-data-revolution-in-us-health-care

Gupta, M. & George, J. F. (2016). Toward the development of a Big Data analytics capability. *Information and Management*, *53*(8), 1049–1064. https://doi.org/10.1016/j.im.2016.07.004

Guttmann, A. (2018). *North America's most in demand skills in data marketing 2017 | Statista*. Retrieved from https://www.statista.com/statistics/818976/north-america-leading-in-demand-skills-data-marketing/

Hashem, I. A. T., Yaqoob, I., Anuar, B., Mokhtar, S., Gani, A., & Khan, U. (2015). The rise of "big data" on cloud computing: review and open research issues. *Information Systems*, *47*, 98–115. https://doi.org/10.1016/j.is.2014.07.006

Homer, M. (2001). Skills and competency management. *Industrial and Commercial Training*, *33*(2), 59–62.

Jafari Navimipour, N. & Zareie, B. (2015). A model for assessing the impact of e-learning systems on employees' satisfaction. *Computers in Human Behavior*, *53*, 475–485. https://doi.org/10.1016/j.chb.2015.07.026

Keser, H. & Semerci, A. (2019). Technology trends, Education 4.0 and beyond. *Contemporary Educational Researches Journal*, *9*(3), 39–49. https://doi.org/10.18844/cerj.v9i3.4269

Khan, N., Yaqoob, I., Hashem, I. A. T., Inayat, Z., Ali, M., Kamaleldin, W., Alam, M., & Gani, A. (2014). Big data: survey, technologies, opportunities, and challenges. *The Scientific World Journal*, *2014*, 1–18.

Klinge, C. M. (2015). A conceptual framework for mentoring in a learning organization. *Adult Learning*, *26*(4), 160–166. https://doi.org/10.1177/1045159515594154

Krasnow Waterman, K. & Bruening, P. J. (2014). Big Data analytics: risks and responsibilities. *International Data Privacy Law*, *4*(2), 89–95.

Laney, D. (2001). 3D data management: controlling data volume, velocity and variety. *META Group Research Note*, *6*(70), 1–4.

Langkafel, P. (Ed.). (2015). *Big data in medical science and healthcare management: diagnosis, therapy, side effects*. Berlin: Walter de Gruyter GmbH & Co KG.

Lee, I. (2017). Big data: dimensions, evolution, impacts, and challenges. *Business Horizons*, *60*(3), 293–303.

Lee, J., Kao, H. A., & Yang, S. (2014). Service innovation and smart analytics for industry 4.0 and a big data environment. *Procedia Cirp*, *16*, 3–8.

Marr, B. (2015). *Big Data: 20 mind-boggling facts everyone must read*. Retrieved from https://www.forbes.com/sites/bernardmarr/2015/09/30/big-data-20-mind-boggling-facts-everyone-must-read/#5947d77c17b1

Miller, S. (2014). Collaborative approaches needed to close the Big Data skills gap. *Journal of Organization Design, 3*(1), 26–30.

Pabedinskaitė, A., Davidavicienė, V., & Milicauskas, P. (2014). Big data driven e-commerce marketing. *Proceedings of 8th International Scientific Conference Business and Management* (pp. 645–654). Vilnius: Vilnius Gediminas Technical University.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., … Duchesnay, E. (2011). Scikit-learn: machine learning in Python. *Journal of Machine Learning Research*, *12*, 2825–2830.

Politaitė, S. & Sabaitytė, J. (2018). Didziuju duomenu naudojimas klientui pazinti. *Mokslas – Lietuvos Ateitis / Science – Future of Lithuania*, *10*, 1–10.

Rajaraman, V. (2016). Big Data analytics. *Resonance*, *21*(8), 695–716. https://doi.org/https://doi.org/10.1007/s12045-016-0376-7

Rao, M. S. (2014). Enhancing employability in engineering and management students through soft skills. *Industrial and Commercial Training*, *46*(1), 42–48. https://doi.org/10.1108/ICT-04-2013-0023

Robinson, M. A., Sparrow, P. R., Clegg, C., & Birdi, K. (2007). Forecasting future competency requirements: a three-phase methodology. *Personnel Review*, *36*(1), 65–90.

Russom, P. (2011). Big data analytics. *TDWI Best Practices Report, Fourth Quarter*, *19*(4), 1–34.

SAP. (2012). Small and midsize companies are looking to make big gains with "big data". Retrieved from http://global.sap.com/corporate-en/news.epx?PressID=19188

Sattayaraksa, T. & Boon-itt, S. (2015). CEO transformational leadership and the new product development process: the mediating roles of organizational learning and innovation culture. *Leadership & Organization Development Journal*, *37*(6), 730–749. https://doi.org/10.1108/LODJ-10-2014-0197

Schoenherr, T. & Speier-Pero, C. (2015). Data science, predictive analytics, and big data in supply chain management: current state and future potential. *Journal of Business Logistics*, *36*(1), 120–132. https://doi.org/10.1111/jbl.12082

Sharma, S. (2015). Rise of Big Data and related issues. *2015 Annual IEEE India Conference (INDICON)* (pp. 1–6). New Delhi: IEEE India Council.

Sivarajah, U., Kamal, M. M., Irani, Z., & Weerakkody, V. (2017). Critical analysis of Big Data challenges and analytical methods. *Journal of Business Research*, *70*, 263–286. https://doi.org/10.1016/j.jbusres.2016.08.001

Smaliukiene, R. & Giedraityte, V. (2018). Research trends in knowledge management in military. In S. Bekesiene & S. Hoskova-Mayerova (Eds.), *Challenges to National Defence in Contemporary Geopolitical Situation (CNDCGS' 2018)* (pp. 142–147). Vilnius: General Jonas Zemaitis Milietary Academy of Lithuania.

Song, J. H., Joo, B. K., & Chermack, T. J. (2009). The dimensions of the learning organization questionnaire (DLOQ): a validation study in a Korean context. *Human Resource Development Quarterly*, *20*(1), 43–64.

Soubra, M., Tanriover, O., & Orhun, A. (2017). Modeling and predicting critical thinking skills of university students using data analytics tool. *Bilisim Teknolojileri Dergisi*, *10*(1), 89–97.

Tanwar, M., Duggal, R., & Khatri, S. K. (2015). Unravelling unstructured data: a wealth of information in big data. *2015 4th International Conference on Reliability, Infocom Technologies and Optimization (ICRITO) (Trends and Future Directions)* (pp. 1–6). Noida: Amity University Uttar Pradesh (AUUP).

Tole, A. A. (2013). Big Data challenges. *Database Systems Journal*, *4*(3), 31–40.

Turkben, A. K., Turkben, E., Karahoca, D., & Karahoca, A. (2016). Big data – a review in health sciences. *Global Journal of Information Technology: Emerging Technologies*, *6*(1), 72–79. https://doi.org/10.18844/gjit.v6i1.392

Tvrdikova, M. (2016). Increasing the business potential of companies by ensuring continuity of the development of their information systems by current information technologies. *Journal of Business Economics and Management*, *17*(3), 475–489. https://doi.org/10.3846/16111699.2013.839475

ur Rehman, M. H., Liew, C. S., Abbas, A., Jayaraman, P. P., Wah, T. Y., & Khan, S. U. (2016). Big data reduction methods: a survey. *Data Science and Engineering*, *1*(4), 265–284.

Waller, M. A. & Fawcett, S. E. (2013). Data science, predictive analytics, and big data: a revolution that will transform supply chain design and management. *Journal of Business Logistics*, *34*(2), 77–84.

Wang, Y. & Li, B. (2015, November). Sentiment analysis for social media images. *2015 IEEE International Conference on Data Mining Workshop* (ICDMW) (pp. 1584–1591). New Delhi: IEEE India Council.

Yin, S. & Kaynak, O. (2015). Big data for modern industry: challenges and trends [point of view]. *Proceedings of the IEEE*, *103*(2), 143–146.