https://doi.org/10.53052/9788367652018.14

Iva KOSTADINOVA[1], Dragan STOJANOVIC[2], Vasyl MARTSENYUK[3], Georgi DIMITROV[1], Dejan RANCIC[2], Oleksiy BYCHKOV[4], Yelyzaveta ZHABSKA[4]

PRZEWODNIK BIZNESOWY PROJEKTU ERASMUS+ IBIGWORLD

Streszczenie: Analityka Big Data pomaga organizacjom używać ich dane i wykorzystywać je do identyfikacji nowych możliwości. To z kolei prowadzi do mądrzejszych ruchów biznesowych, bardziej wydajnych operacji, wyższych zysków i zadowolonych klientów. Firmy korzystające z Big Data z zaawansowaną analityką zyskują na wielu sposobach. Firmy muszą być edukowane na temat możliwości, jakie daje Big Data i jak rozwijać się poprzez analizę Big Data. W wyniku realizacji projektu Erasmus+ „Innowacje dla Big Data w Realnym Świecie" (iBIGworld) 2020-1-PL01-KA203-082197 powstało szkolenie. W pracy przedstawiono tematy, które są planowane na szkolenia w organizowanym kursie oraz w których z nich planowane jest włączenie przedstawicieli biznesu. Zaznacza się również, w jakiej formie planowane jest zaangażowanie przedstawicieli biznesu w szkolenie.

Słowa kluczowe: biznes, przewodnik, szkolenie Big Data, iBIGworld

BUSINESS GUIDE FOR IBIGWORLD ERASMUS+ PROJECT

Abstract: Big data analytics helps organizations harness their data and use it to identify new opportunities. That, in turn, leads to smarter business moves, more efficient operations, higher profits and happier customers. Businesses that use Big Data with advanced analytics gain value in many ways. Businesses need to be educated about the opportunities provided by Big Data and how to develop by analysing Big Data. As a result of the Erasmus + project "Innovations for Big Data in a Real World" (iBIGworld) 2020-1-PL01-KA203-082197 the training course has been developed. In the work there are presented the topics that are planned for training in the organized course and in which of them there is a planned inclusion of representatives from the business. It is also is noted in what form it is planned to involve the representatives of the business into the training course.

Keywords: business, guide, Big Data training, iBIGworld

[1] University of Library Studies and Information Technologies (ULSIT), Sofia, Bulgaria: (i.kostadinova, g.dimitrov)@unibit.bg
[2] University of Niš (UNi), Nis, Serbia: (dragan.stojanovic, dejan.rancic)@elfak.ni.ac.rs
[3] Department of Computer Science and Automatics, University of Bielsko-Biala, Poland: vmartsenyuk@ath.bielsko.pl
[4] Taras Shevchenko National University of Kyiv (TSNUK), Kiev, Ukraine: oleksiibychkov@knu.ua, y.zhabska@gmail.com

1. Introduction

Big Data and its analysis techniques are at the center of modern science and business. The size and number of available data sets is growing rapidly with the collection of data from various devices, such as mobile, low-cost and numerous IoT devices, software logs, cameras, microphones, RFID readers and wireless sensor networks.
The world's technological capacity per capita to store information has doubled every 40 months since the 1980s. The data can be seen in Fig. 1 and Fig. 2.

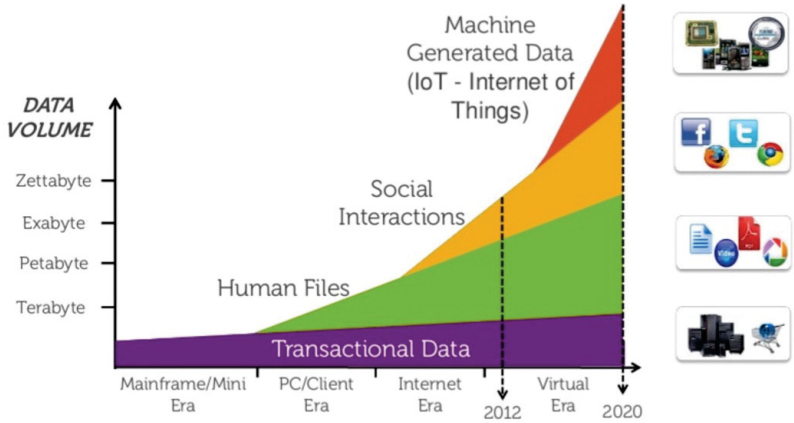

Figure 1. Growth and digitalization of the global data volume to 2020 by type

Relational database management systems and desktop statistical software are very difficult to process and analyse Big Data. Such tasks usually require software that runs in parallel on tens, hundreds or even thousands of servers.
Big data often includes data that exceeds the capacity of traditional software to process it in a timely manner. Challenges to analysing large data sets include data capture, storage, analysis, sharing, transfer, visualization, retrieval, updating and protection.
Through Big Data analysis, analyses of processes such as consumer preferences, analyses of the quality of a product placed on the market, analyses and forecasts for the development of companies and many other applications can be made.

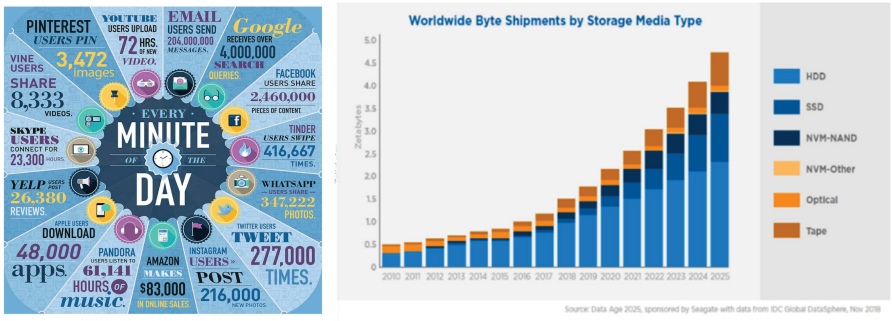

Figure 2. Growth and digitalization of the global storage capacity to 2025 by type storage

The implementation of Big Data solutions in various areas is related with the development of the innovative training courses [1-5]. The given work is the result of Erasmus + Project "Innovations for Big Data in a Real World" (iBIGworld) 2020-1-PL01-KA203-082197 [6]. The aim of the project was to conduct a study in relation to the use and operation of Big Data and the need for the labor market for such specialists. The following guideline was developed as a result of the short-time learning activities:

- Training for teachers,
- Training for students,
- Training for business representatives.

The organized training was in the form of a summer school - a seminar, where teachers and students met to exchange experience and new knowledge.

2. Programme's essence

iBIGWorld was created like an experiential learning programme which, in addition to taught concepts related to the Big Data, engages students in challenges and problems posed by a companies. Its aim is to enhance collaboration between business and universities by developing an innovative approach and methodology for teaching Big Data topics.

The total programme consists of 120 hours during which there will be 12 weekly sessions (topics) of 10 hours each (4 hours lectures + 6 hours learning activities). Business representatives got involved in teaching students to work with Big Data, stating their problems. In this way, they provided real problem situations for the students to analysed and find a solution. Additional work on practical problems related to the Big Data presented by Business Managers (BM) is foreseen during the learning sessions.

3. Business managers participation

Figure 3 shows the learning organization and topics in the Big Data course.

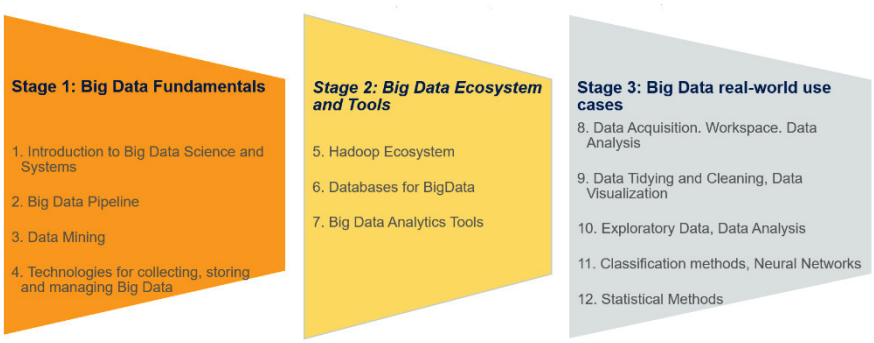

Figure 3. Training course strukture

In Table 1, you can see what topics are planned for training in the organized course and in which of them there is a planned inclusion of representatives from the business. It is also was noted in what form it is planned to involve the representatives of the business.

Table 1. Course topic and business manager presence

No.	**Topic**	**Business managers presence**
1.	Introduction to Big Data science and systems	
2.	Big Data Pipeline	Presenting the practical problems
3.	Data Mining	
4.	Technologies for collecting, storing and managing Big Data	Presenting technical requirements
5.	Hadoop ecosystem	
6.	Databases for Big Data	Discussion of the ideas
7.	Big Data analytics tools	
8.	Data acquisition, workspace, data analysis	Discussion of the ideas
9.	Data tidying and cleaning, data visualization	
10.	Exploratory data, data analysis	Presentations of draft solutions
11.	Classification methods, Neural Networks	
12.	Statistical methods	Presentations of final solutions

As a final, in a workshop, the business managers participated in presentation a follows theme:

- Approval of the Elaborated Big Data Requirements,
- Big Data Requirements and how to find the best solution to problems,
- Big Data specialists in the Data Lake ecosystem,
- iBigData framework for training in HE,
- Big Data Smart JobHub.

4. Business managers (BM) engagement

BMs would optimally attend the programme on 6 occasions for a maximum of 2 hours to meet with the students. BMs would also be expected to additionally devote one hour per week to communicate with the Mentors through MS Teams platform in order to address any student queries or questions that may arise out of the scheduled meetings.

5. Meetings with Business Managers

Meeting 1:
- Provide background material about the company.
- Outline the challenge/business Big Data problem.
- Detail consequences of not finding a solution.

Meeting 2:
- Identify any known constraints/limiting factors to resolving the challenge/business Big Data problem.
- Discuss technology platforms for students' solutions.

Meetings 3, 4:
- Discuss the progress of the students' work.
- Make corrections if needed.
- Provide clarifications/additional information if needed.

Meeting 5:
- Students present draft solutions.
- Business Manager to provide feedback to students.

Meeting 6:
- Business Manager to select most suitable solution for their business.
- Official presentation by the student teams of their results/proposed concepts.

6. Business challenges

- Business challenges represent some of the problems and challenges the companies face and offer to members of the scientific research community to find solutions for these problems.
- At the end of the programme the preferred solution will be selected by BMs.
- Business challenges aims to ensure that companies get the integration of internal and external knowledge and resources into development of the new Big Data system and architecture.
- The cooperation with the private sector aims to solve technical, technological, organizational and other Big Data problems, shifts traditional research towards practical applications, while encouraging innovation and creativity of students.

6.1. Business challenge - the problem

- Should not be too complex for students' abilities.
- Must be a real and genuine issue that is affecting the organization's viability and/or growth.
- It must be relevant to the Big Data area.
- The BMs must not have previously solved the business problem.

6.2. Business challenge form

- The information for the challenge needs to cover two separate parts, the one related to the company information and the other dedicated to challenge itself.
- The part with Challenge needs information about Challenge name, Short challenge description, Consequences of not resolving this problem, and Specific requirements to be met.
- This form will be given to business representatives, and they should fill in with real information based on the instructions provided below.

6.3. Big Data challenge – company info

Table 2 presents the first business challenge in the Big Data course

Table 2. Course topic and business manager presence

Company name:	**City Heating Plant Nis**
Short company description (max 300 words)	Public utility company City Heating Plant is a company for the production and distribution of thermal energy on the territory of the city of Nis
Contact person information	Firstname Lastname example@gmail.com +38118 xxx xxx
Challenge name: **Short challenge description (500 words):**	Predictive maintenance of thermal substations
Thermal	Substation serves for the regulated exchange of thermal energy between the primary heating network of the city and the secondary network of radiator heating subsystem of individual residential buildings…
Consequences of not resolving this problem (200 words):	Valve failure leads to damage to the entire substation and consequently great costs of replacing the substation and interruption to the thermal energy delivery to citizens…
Specific requirements to be met:	The dedicated budget for the solution is around 5000 euros...

6.4. Others Big Data challenges, presented in the training course

Other challenges in the Big Data course are presented in Table 3.

Table 3. General description of the presented use cases

	Use case (link to dataset)	branch	types of data used
1	COVID-19 (numeric, labels) *https://github.com/covid19datahub/R*	public health	numeric, characters, and labels
2	Historical Daily Weather Data 2020 *https://www.kaggle.com/vishalvjoseph/we ather-dataset-for-covid19-predictions*	weather	numeric, characters, and labels
3	Global Health Data Exchange *http://ghdx.healthdata.org/*	public health	numeric, characters, and labels
4	Sensors of air pollution (geolocational data)	ecology	numeric, characters, and labels
5	Vessels segmentation (images) *https://www.idiap.ch/software/bob/docs/b ob/bob.db.drive/stable/index.html*	medicine	images
6	Google Street View House Number (SVHN) / Dataset (images) *https://github.com/aditya9211/SVHN-CNN*	navigation	images
7	Physionet (signals) (physionet.org)	medicine	signal
8	Natural Language Processing (texts) and application for Stocks	literature	text
9	Stock exchange data *https://www.kaggle.com/mattiuzc/stockexchange-data*	finance	numeric, characters, and labels
10	Biochemical reaction study	science	signal

Conclusion

Joining business representatives in Big Data training has proven fruitful. Business representatives saw how they could use Big Data to the advantage of their companies. They posed problems they faced and gave the students a chance to release with a solution for them.
Collaboration between business and learning process is mutually beneficial for both sides - learning and business environment.

REFERENCES:

1. DUBEY RAMESHWAR, ANGAPPA GUNASEKARAN: Education and training for successful career in Big Data and business analytics. Industrial and Commercial Training 47.4 (2015), 174-181.
2. MARCIANO R., et al.: Archival records and training in the age of Big Data. Re-Envisioning the MLS: Perspectives on the future of library and information science education. Emerald Publishing Limited, 2018.
3. RUIZ-PALMERO J., et al.: Big data in education: Perception of training advisors on its use in the educational system. social sciences 9.4(2020)53.

4. GREENE A. C., et al.: Adapting bioinformatics curricula for Big Data. Briefings in Bioinformatics 17.1 (2016), 43-50.
5. DANIEL B.: Big Data and analytics in higher education: Opportunities and challenges. British journal of educational technology 46.5 (2015): 904-920.
6. iBigWorld: Innovations for Big Data in a Real World (Erasmus+ project 2020-1-PL01-KA203-082197) documentation *https://ibigworld.ath.edu.pl/index.php/en/home-english/*