DOI: https://doi.org/10.53052/9788367652018.16

Georgi DIMITROV[1], Iva KOSTADINOVA[1], Vasyl MARTSENYUK[2], Dejan RANCIC[3], Oleksiy BYCHKOV[4], Eugenia KOVATCHEVA[1], Aleksandra KŁOS-WITKOWSKA[2]

SUKCESY I OSIĄGNIĘCIA PROJEKTU iBIGWORLD

Streszczenie: Niniejsza praca przedstawia sukcesy i osiągnięcia projektu Erasmus+ nr. 2020-1-PL01-KA203-082197 pt. „Innowacje dla Big Data w realnym świecie" (iBIGworld). Wyniki projektu pokazują kompleksowy wpływ programu Erasmus+ na poprawę jakości kształcenia i powiązanie programów uniwersyteckich z biznesem.

Słowa kluczowe: Big Data, umiejętności twarde, umiejętności miękkie, iBIGworld

iBIGWORLD PROJECT SUCCESSES AND ACHIEVEMENTS

Abstract: This work presents the successes and achievement of the Erasmus+ project no. 2020-1-PL01-KA203-082197 entitled "Innovations for Big Data in a Real World" (iBIGworld). The results of the project show the comprehensive impact of the Erasmus+ program on improving the quality of education and linking university curricula with business.

Keywords: Big Data, hard skills, soft skills, iBIGworld

1. Introduction

The global data was started to grow exponentially in the 21st century, it has shown no signs of slowing down. They accumulate mainly via the internet, including social networks, web search requests, text messages, media files, IoT devices, and sensors [1-8]. Now, the world is powered by Big Data. Companies are forced to seek experts in Big Data for consultations, and capabilities to harness complex data processing.
The big question is how to harness the power that Big Data gives us!
The most significant obstacle preventing organizations from realizing the full potential of their data assets is widespread data disorder. Governments and private companies have quickly accumulated big massive of data, and adopted the environments to store them. And while insights might be buried within all that raw

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

data; if no one knows where it came from, how to find it, what it means or if they can trust it, it will remain untapped and untouched.
All organizations have a pressing need for professionals who understand Big Data and have the competence and knowledge to manage it.
The challenge faced in light of the data revolution is finding people with the required set of skills to transform data into actionable insight. The interest on the role of the data scientist and the related data analytics skills is growing.
Large amounts of data are generated and analyzed by various Big Data professionals. At the same time, the demand for Big Data professionals with a competencies in this areas increases. On the other hand, the supply remains low which creates great job opportunities for individuals within this field.
Higher Education Institutes (HEIs) unable to cope with this fast transformation due to the rapid pace of changing and in some cases due to less flexibility in learning modes. This mismatch becomes even more intensive in Big Data, an area where trends raise urgent challenges and needs of high-qualified personnel.
As a result, an inefficient loop is observed: centers of excellence struggle to find talented and skilled young specialists while graduated students are in lack of practical skills required in labor market and industry. According NIST and SAS Big Data Analytics: “Big data analytics examines large amounts of data to uncover hidden patterns, correlations and other insights. With today’s technology, it’s possible to analyze your data and get answers from it almost immediately – an effort that’s slower and less efficient with more traditional business intelligence solutions.”
In order to try to solve some of these problems, the project no. 2020-1-PL01-KA203-082197 entitled “Innovations for Big Data in a Real World” – iBigWorld [9] showed under the auspices of the Polish ERASMUS+ agency brought together partners from European universities – University of Bielsko-Biala (Poland), University of Library Studies and Information Technologies (Bulgaria), University of Nis (Serbia) and Taras Shevchenko National University of Kyiv (Ukraine). This collaboration has provided innovative solutions to develop Big Data experts. The training framework is based on the IEEE Guidelines for Big Data in Machine Learning.

2. Main part

Facing the successful use of Big Data in the real world can be successfully resolved managed to solve the following problems:

a. Take the first steps towards bridging the current gap in digital skills in the project partner countries, building on the European Commission's 'e-Skills for jobs' campaign".
b. Build an ecosystem of key partners to create a gateway to underrepresented talent pools.
c. Identifies underrepresented skills, the rationale behind the phenomenon of talented people lacking the traditional qualities to find good work, and Big Data with the most pressing needs.
d. With the help of the created Smart Big Data Job Hub, it was able to encourage business to use the results of the project, through active cooperation between universities and business, including the creation of new internship programs based on Big Data, supervision of PhD students between higher education institutions and business.

e. Implement a range of new initiatives ensuring that these skills are adequately promoted in the curriculum, teacher development, assessment practices and learning content.
f. Formulate trends in ICT, Big Data and AI, create new educational tools and learning resources that are collected, processed and distributed through the Smart Big Data Job Hub for curriculum modernization and beyond.

3. Partners

Project "Innovations for Big Data in a Real World" (iBIGWorld) joined teachers, and students from 4 Balkans Universities and business organizations, as follows:

- University of Bielsko-Biala (Poland),
- University of Library Studies and Information Technologies (Bulgaria),
- University of Nis (Serbia),
- Taras Shevchenko National University of Kyiv (Ukraine).

Table 1 shows how many participants from each of the universities are in the project iBIGWorld.

Table 1. Participants from each University in iBIGWorld

University	**Participants from a University**
University of Bielsko-Biala (Poland)	5
University of Library Studies and Information Technologies (Bulgaria)	24
University of Nis (Serbia)	11
Taras Shevchenko National University of Kyiv (Ukraine).	8
Total:	48

4. Project stages

The project was organized in 4 stages, spread over a period of 2 calendar years.

4.1. Stage 1. Survey

In the first phase of project 2020-1-PL01-KA203-082197 "Innovations for Big Data in a Real World" (iBIGworld) under the Erasmus+ program, a study was conducted on the state of the subject area, namely Big Data in a Real World and innovations in it. This project aims to join together Universities, businesses and provide innovative solutions to develop Big Data experts. The survey was performed online using google forms tools.

Due to various formats and specifications of the found information in each case, the data was collected by scientists based on phrase search. Several search phrases were used: "Big Data", "good practice" and "specification". The survey was performed during a period from the 1'th of September 2020 to the 28'th of February 2021. To obtain a wide range of data multiple question fields, with an additional open-field option, were offered to mitigate the effect of narrowed answers suggestions.

4.1.1. Areas of researching

The survey contains 7 online forms with both open and closed questions (Table 2). The questions consider different questions for job offerings, labor market in the field of Big Data, existing training programs, and good practices, and collecting IT specifications of good practices in Big Data. Also for IT graduates Masters and IT Alumni in Information Systems and Technologies and for Employers to Specifying Graduate Competencies.
To make the process of data collection unbiased no additional recommendation was added. No events were reported during that time that could influence the result.
Stage 1 was realized through several on-line forms in different directions and using different target gropus, namely:

Table 2. Reasearching in stage 1

No	Objective of the research in the field of Big Data	Source of data
Research 1:	Existing training programs	Offers on websites
Research 2:	Demands of labor market	Offers on job portals
Research 3:	Existing scientific projects	Web-portals of the scientific programs and projects
Research 4:	Competencies and skills for teachers	Surveys for academic lecturers
Research 5:	Competencies and skills for trainees	Surveys for IT graduates Masters and IT Alumni in Information Systems and Technologies
Research 6:	Graduate competencies	Questionnaire for Employers
Research 7:	Good practice in Big Data	Collecting IT specifications

4.1.2. Collection and analysis of data

The research was conducted by scientists from the 4 countries - participants in the project - Poland, Ukraine, Bulgaria Serbia.
The surveys were made without the numbering of the questions. This final report contains the conclusions for each report of surveys.

4.2. Stage 2. Setup the Big Data requirements

This stage was aimed at identifying soft and professional competencies, learning outcomes that shape competencies, and educational topics that shape learning outcomes. According to the results of the survey was prepared a frame, that can be used to form educational programs for Big Data courses.

4.2.1. Syllabus development methodology

A survey of stakeholders in seven nominations was conducted. According to the results of the survey, general and professional competencies were identified. which must be formed in students in the learning process. The competencies that received the most votes of the surveyed stakeholders were suggested for consideration when compiling the syllabus.
From the table of the soft skills, the competencies rated from 431 to 200 were selected for inclusion in the curricula.

Moreover, there were developed the table presenting the professional competencies proposed to be included in the curricula. When selecting professional competencies, we were guided primarily by the requirements of employers and the labour market. The requirements of teachers and graduates were taken into account to develop the motivation of teachers and students as participants in the educational process.

4.2.2. Results

Based on the determined importance of soft and hard competencies, 2 training frameworks were developed:
Study topics for the formation of soft skills - presents soft skills competencies, learning outcomes, and course topics that form the relevant competencies,
Study topics for the formation of professional skills - presents professional competencies, learning outcomes, and course topics that form the relevant competencies.
A matrix of competences and topics has been prepared.
The following documents were used to determine learning topics and learning outcomes:

1. E-Competence Framework
2. ACM DataScience_Final_Report2021.pdf
3. ACM Computing Competencies for Undergraduate Data Science Curricula 2019.pdf

4.3. Stage 3. Forming structure of training course and training materials

The 120-hour training course is broken down into 3 stages contain 12 Units. Namely, we can describe the course in the following way

- 12 topics,
- 10 hours per topic,
- 4 hours for the lectures per topic,
- 6 hours of learning activities per topic.

4.3.1. Defining the training themes and grouping in the modules

The training course is organized into 3 stages with several themes in every stage. (Fig. 1).

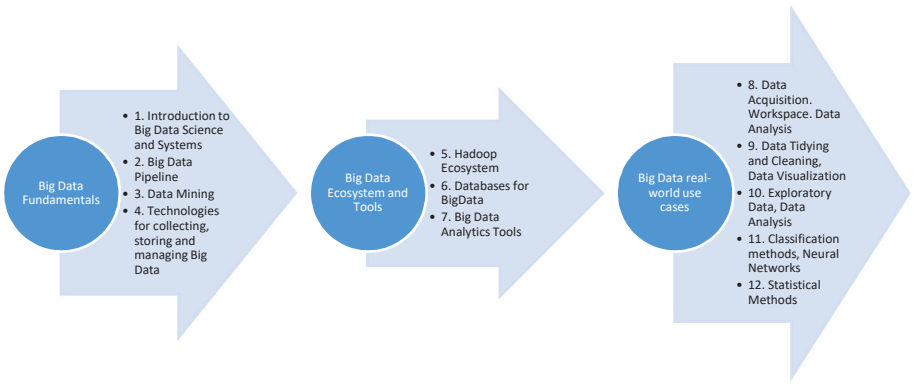

Figure 1. Training course structure

4.3.2. Defining the training methods. Training course details (activities)

In the the course are using several learning methods, the basis of which is self-learning. Each of the unit contains following activities:

- Lectures (presentations and lecture notes),
- Demonstration of real world use cases (presentations),
- Practical tasks in a team (practical exercises),
- Q&A Sessions (questions and discussion),
- Learning Scenarios (branching scenarios).

Rather than academic theory, the course content is practical and based in the real world Big Data use cases. The course is developed with student participation, expert facilitation, and the application of new approaches and techniques in mind. The training program is designed to match the skillsets of the most successful practitioners.

4.3.3. Defining the training techniques

The following strategies should be included in the instructional methods. The trainers are not limited to these tools and approaches, and they are encouraged to try out they own concepts:

- Energisers, icebreakers, and games,
- Seminars,
- Self-directed education,
- Self-study and group conversations,
- Individual and group contemplation, as well as experience sharing,
- Research studies,
- Role-playing exercises.

4.3.4. Defining structure of themes of Big Data Course

In Figure 2 there is shown the structure that is assumed for each of the topics covered in the Big Data training.

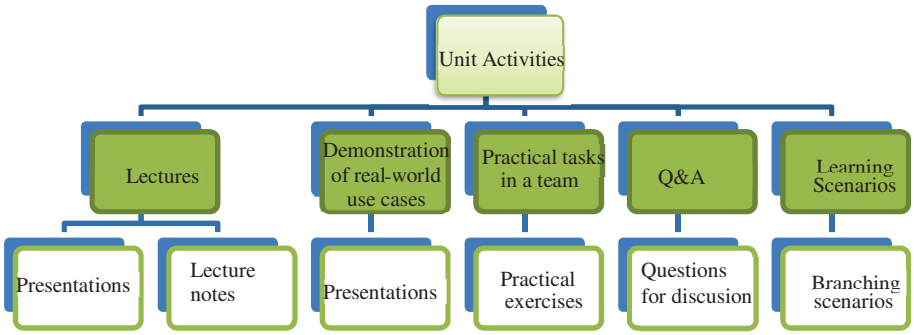

Figure 2. Unit Activities

4.3.5. Preparing the lectures, presentations, questions for discussion, and quizzes for each theme of Big Data Course

The preparation of the lecture material was distributed equally among the participants - partners in the project. Each of the universities undertook to prepare 3 topics from the curriculum.

4.4. Stage 4. Training course and training

As a result of an Erasmus programme "iBigWorld: Innovations for Big Data in a Real World - competence-based learning in Big Data", a training was created. 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. In line with the project, the following were organized:

- 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. The total programme consists of 120 hours learning during which there were 12 weekly sessions (topics) of 10 hours each (4 hours lectures + 6 hours learning activities).

PARTICIPANTS:
StudentsŁ

- Four students from UBB, Poland,
- Four students from ULSIT, Bulgaria,
- Four students from UNi, Serbia.

TrainersŁ

- Four trainers from UBB, Poland,
- Four trainers from ULSIT, Bulgaria,
- Four trainers from Uni, Serbia,
- Three trainers from TSUNK, Ukraine (online).

Business organisationŁ

- Four bussines organisations from Serbia,
- Two bussines organisations from Bulgaria.

ORGANIZATION

- Theoretical Background: Lectures, Individual Research Tasks, Quizzes,
- Practical Sessions: Live and Video Demonstrations, Guidelines, Labs, Individual and Group Tasks,
- Assessment: Final Group Project.

LEARNING METHODS

- Learning-by-doing,
- Real-Case and Project-based Learning,

Additional work on practical problems related to the Big Data presented by Business Managers (BM) is foreseen during the learning sessions. An overview of the

programme is shown below with times when BM would be required to brief the students, deal with ongoing queries and receive presentations:
Live meeting between BM and trainers and students was recommended but not required. BM attended in the programme on 6 occasions for a maximum of 2 hours to meet with the students. The BM devoted additional time per week to communicate with the Mentors through MS Teams platform in order to address student queries or questions.

5. Multiplier Event

Multiplier Event E1 has been an integral part of the iBIGworld project's dissemination and exploitation of results activities. Planned to take place in June 2022, the events took place successful in June in Poland, Bulgaria and Serbia and in September in Ukraine. More than 280 participants participated in the four events on site and about 200 online (over 480 in total).

6. Final results and outputs

The primary aim of the iBIGWorld programme is to enhance collaboration between businesses and universities by developing an innovative approach and methodology for teaching Big Data content. It is intended that the programme will become part of the university curriculum as a course with relevant credits. iBIGWorld is an experiential learning programme which in addition to taught concepts related to the Big Data, engages students in challenges and problems posed by a real business.
iBIGWorld foresees the development new curriculum, related to the use of Big Data to enhance the quality of management of modern ecosystems. In this way, the students will be enhancing their knowledge through the use of Big Data IT tools and by sharing aspects of training on Big Data management skill issues, and by doing so improving accessibility for everyone.
Many people were involved in the development of the project:

- a total of 48 representatives of the four Universities, participants in the project,
- more than 80 academic lecturers from other universities who participated in the study.
- 631 students from the four countries - participants in the project, whose education is in the field of IT graduates Masters and IT Alumni in Information Systems and Technologies,
- More than 80 representatives from the business world.

Training materials for a Big Data course were prepared (lectures, presentation, case studies, question for discussion, and quizzes).

An e-course on Big Data was organized with study materials in English. In parallel with this, electronic courses on Big Data were created in the languages of each of the Universities participating in the project - Polish, Serbian, Bulgarian and Ukrainian.

7. Implementation of the initially set indicators for the project

The implementation of the project can be presented with the help of quantitative indicators in Table 3

Table 3. The quentative indicators of the project

Indicators	Realized indicators
Number of good practices analyses (at least 10)	15
One curriculum developed addressing the skillgaps identified	1
One methodology to anchor the program into the local environments (Industry, HEIs, local organizations) securing fruitful cooperation and direct interconnections among all stakeholders	1
One skills Foresight Study appointing concrete skills gaps and talent needs	1
One common BIG DATA platform/smart Big Data job hub engine developed	1
Number of staff trained at the transnational level (at least 20)	21
Number of students involved (at least 20 from partners organizations and 20 from other HEIs networks)	631
Number of academics involved (at least 4 from partners organizations and 60 from other HEIs networks)	80
Number of entrepreneurs involved (at least 4 from partners organizations and 20 from partners networks)	38
Number of regional stakeholders involved in the pilot (from academy, industry, government) (at least 1 representative for each target group per partner country)	12
Number of communication materials developed (website, leaflets, mailing lists)	website – 3, leaflets – 400, mailing lists - 6
Number of press appearances discussing the related skill alliances mechanisms	62
Number of online science facilitation resource accessed through IBIGWORLD open smart Big Data job hub engine.	124765
Number of online user return rates to IBIGWORLD online presence means	1467
Number of HEIs reached by IBIGWORLD communication (at least 30)	More than 80

8. Conclusions

The Big Data Innovations in the Real World (iBIGworld) project really brought academia and business together. This collaboration has provided innovative solutions to develop Big Data experts.

The results show that bridging of the current digital skills gap in the project countries has begun. An ecosystem of key partners has been built to create a gateway to underrepresented talent pools.

Underrepresented skills among students have been identified. which actually helps them find a good job.

Thanks to the Smart Big Data Job Hub, contacts between universities and businesses have been promoted, including new internship programs based on Big Data, supervision of PhD students between higher education institutions and businesses.

The new course created has really helped universities to create new courses with a trend towards new curricula in ICT, Big Data and AI by adding new educational tools and learning resources.

All this shows that the initial tasks set before the project partners were fulfilled, which is evidenced by the implementation of the initially set indicators.

The project showed that the participation of universities in Erasmus+ programs significantly expands the opportunities of young specialists to realize themselves in the field of high technologies.

REFERENCES

1. BELLANDI V.: A Big Data Infrastructure in Support of Healthy and Independent Living: A Real Case Application (2023) Intelligent Systems Reference Library, 229, pp. 95-134. DOI: 10.1007/978-3-031-11170-9_5.
2. ARIFIN S., SILALAHI F.E.S., PRAYITNO M., MAJID N.K., AMHAR F., GULARSO H.: Geospatial Big Data Management Testing Using Open Source Technology (2023) Mechanisms and Machine Science, 121, pp. 29-42. DOI: 10.1007/978-3-031-09909-0_3.
3. ZHANG S., OU W., REN G., WANG H., ZHU P., ZHANG W.: Risk Model and Decision Support System of State Grid Operation Management Based on Big Data (2023) Lecture Notes on Data Engineering and Communications Technologies, 122, pp. 419-427. DOI: 10.1007/978-981-19-3632-6_51.
4. JOSE D.T., HOLME J., CHAKRAVORTY A., RONG C.: Integrating Big Data and blockchain to manage energy smart grids—TOTEM framework (2022) Blockchain: Research and Applications, 3 (3), art. no. 100081, . DOI: 10.1016/j.bcra.2022.100081.
5. ESPOSITO S., ORLANDI S., MAGNACCA S., DE CURTIS A., GIALLUISI A., IACOVIELLO L.: on behalf of The Neuromed Clinical Network Big Data and Personalised Health Investigators, Clinical Network for Big Data and Personalized Health: Study Protocol and Preliminary Results (2022) International Journal of Environmental Research and Public Health, 19 (11), art. no. 6365. DOI: 10.3390/ijerph19116365.
6. ZHANG G.PyCLKDE: A Big Data-enabled high-performance computational framework for species habitat suitability modeling and mapping (2022) Transactions in GIS, 26 (4), pp. 1754-1774. DOI: 10.1111/tgis.12901.
7. DEEPA N., PHAM Q.-V., NGUYEN D.C., BHATTACHARYA S., PRABADEVI B., GADEKALLU T.R., MADDIKUNTA P.K.R., FANG F., PATHIRANA P.N.: A survey on blockchain for Big Data: Approaches, opportunities, and future directions (2022) Future Generation Computer Systems, 131, pp. 209-226. DOI: 10.1016/j.future.2022.01.017.
8. KASTOUNI M.Z., AIT LAHCEN A.: Big data analytics in telecommunications: Governance, architecture and use cases (2022) Journal of King Saud University - Computer and Information Sciences, 34 (6), pp. 2758-2770. DOI: 10.1016/j.jksuci.2020.11.024.
9. iBigWorld: Innovations for Big Data in a Real World (Erasmus+ project 20201PL01KA203082197) documentation *https://ibigworld.ath.edu.pl/index.php/en/homeenglish/*