COURSE DESCRIPTION CARD - SYLLABUS

Course name

Big Data in Management

Course

Field of study

Logistics

Year/Semester

3/5

Area of study (specialization)

Profile of study

general academic

Level of study

First-cycle studies

Course offered in

English

Form of study

part-time

Requirements

compulsory

Number of hours

Lecture

8

Laboratory classes

8

Other (e.g. online)

Tutorials

8

Projects/seminars

Number of credit points

4

Lecturers

Responsible for the course/lecturer:

Ph.D., Eng. Michał Trziszka

Mail to: michal.trziszka@put.poznan.pl

Faculty of Engineering Management

ul. J. Rychlewskiego 2, 60-965 Poznań

Responsible for the course/lecturer:

Prerequisites

Knowledge of relational database systems. Knowledge of the SQL language. Basic knowledge of programming languages, eg C #.

Course objective

Provide students with basic knowledge in the field of organization, management and processing of Big Data (very large data sets). Developing students' ability to solve problems related to the organization, management and processing of Big Data.

Course-related learning outcomes

Knowledge

1. Student knows the basic concepts of logistics and its detailed issues and supply chain management With Big Data support. [P6S_WG_05]

2. Student knows the best practices in logistics and its specific issues using Big Data. [P6S_WK_06]

Skills

1. Student can design, using appropriate methods and techniques, an object, system or process that meets the requirements of logistics and its specific issues and supply chain management with Big Data usage in enterprise. [P6S_UW_07]

2. Student is able to present, using properly selected means, the problem within the scope of logistics and its specific issues, and supply chain management using Big Data. [P6S_UK_01]

Social competences

1. Student is able to plan and manage in an entrepreneurial manner in Big Data supporting company. [P6S_KO_01]

2. Student is aware of cooperation and group work on solving problems within logistics and supply chain management with Big Data usage in company. [P6S_KR_02]

3. Student is aware of the responsible filling, correct identification and resolution of dilemmas related to the profession of logistics with usage of Big Data. [P6S_KR_01]

Methods for verifying learning outcomes and assessment criteria

Learning outcomes presented above are verified as follows:

Lecture: The knowledge gained during the lecture is verified by one test at the last lecture. The test consists of 10-15 questions (test and open-ended), with different scores. Passing threshold: 50% of points. Final issues, on the basis of which the questions are prepared, will be sent to students by e-mail using the university's e-mail system.

Tutorial: During the exercises, students work in groups on specific topics that they present in the form of a multimedia presentation. For each of the 7 tasks, students receive grades (7 grades). The final grade is the average of the 7 grades. The content of the tasks is related to the topic, and the scope of tasks includes lecture issues.

Laboratory: During the laboratory, the laboratory on the system over the available IT topics. The results are at the end of each class. For each of the 7 statements, a guaranteed evaluation rating (7 ratings). The final grade is from these 7 grades. The content of the statements is linked to the topics as well as the information titles about expanding the lectures.

Programme content

Lecture: Introduction to Big Data systems, motivations, definitions, problems in the world of Big Data. Hadoop platform, distributed file systems on the example of HDFS, task scheduling systems in Big Data systems on the example of YARN, data batch processing engines on the example of MapReduce, MapReduce processing optimization techniques, decomposition of complex problems into MapReduce action sequences, Hadoop Streaming. Relational data processing using Spark SQL, DataFrame and Dataset data types, data processing in Spark SQL, processing optimization mechanisms. The use of Big Data in Logistics.

Tutorial: Introduction to Big Data systems, motivations, definitions, problems in the world of Big Data. Hadoop platform, distributed file systems on the example of HDFS, task scheduling systems in Big Data systems on the example of YARN, data batch processing engines on the example of MapReduce, MapReduce processing optimization techniques, decomposition of complex problems into MapReduce action sequences, Hadoop Streaming. Relational data processing using Spark SQL, DataFrame and Dataset data types, data processing in Spark SQL, processing optimization mechanisms. The use of Big Data in Logistics.

Laboratory: Introduction to Big Data systems, motivations, definitions, problems in the world of Big Data. Hadoop platform, distributed file systems on the example of HDFS, task scheduling systems in Big Data systems on the example of YARN, data batch processing engines on the example of MapReduce, MapReduce processing optimization techniques, decomposition of complex problems into MapReduce action sequences, Hadoop Streaming. Relational data processing using Spark SQL, DataFrame and Dataset data types, data processing in Spark SQL, processing optimization mechanisms. The use of Big Data in Logistics.

Teaching methods

Lecture: multimedia presentation illustrated with examples given on the board, discussion and problem analysis.

Tutorial: problem solving, discussion, team work.

Laboratory: problem solving, discussion, team work.

Bibliography

Basic

1. Marz N., Warren J., Big Data. Principles and best practices of scalable realtime data systems, Manning Pubications Co., 2015.

2. White T., Hadoop. Kompletny przewodnik. Analiza i przechowywanie danych, Helion, Gliwice, 2015.

3. Zaharia M., Chambers B., Spark: The Definitive Guide, O'Reilly Media, 2018.

4. Odersky M., Spoon L., Venners B., Programming in Scala, 3rd edition, Artima Inc, 2016.

5. Rajaraman A., Ullman J.D., Mining of Massive Datasets, Cambridge University Press, 2012 (http://infolab.stanford.edu/~ullman/mmds.html).

6. Garcia-Molina H., Ullman J.D., Widom J., Systemy baz danych. Kompletny podręcznik, Helion, Gliwice, 2015.

Additional

1. Ryza S., Lasersson U., Owen S., Wills J., Spark. Zaawansowana analiza danych, Helion, Gliwice, 2015.

2. Horstmann C., Scala for the Impatient, Addison-Wesley, 2016.

3. Królikowski Z., Hurtownie danych: logiczne i fizyczne struktury danych, Wydawnictwo Politechniki Poznańskiej, Poznań, 2007.

Breakdown of average student's workload

	Hours	ECTS
Total workload	100	4,0
Classes requiring direct contact with the teacher	24	1,5
Student's own work (literature studies, preparation for laboratory classes/tutorials, preparation for tests, project preparation) [1]	76	2,5

[1] delete or add other activities as appropriate