FACULTY OF PURE AND APPLIED MATHEMATICS
SUBJECT CARD

Name in Polish WSTĘP DO ANALIZY DUŻYCH WOLUMENÓW DANYCH
Name in English INTRODUCTION TO BIG DATA ANALYTICS
Main field of study (if applicable): APPLIED MATHEMATICS
Specialization (if applicable): DATA ENGINEERING
Level and form of studies: ~~1st~~/ 2nd* level, ~~uniform magister studies*~~, full-time / ~~part-time*~~
Kind of subject: ~~obligatory~~ / optional ~~/ university-wide*~~
Subject code
Group of courses YES ~~/ NO~~*

	Lecture	Classes	Laboratory	Project	Seminar
Number of hours of organized classes in University (ZZU)	30		30		
Number of hours of total student workload (CNPS)	90		60		
Form of crediting	crediting with grade				
For group of courses mark (X) final course	X				
Number of ECTS points	3		2		
including number of ECTS points for practical classes (P)			4		
including number of ECTS points corresponding to classes that require direct participation of lecturers and other academics (BU)	3				

*delete as applicable

PREREQUISITES RELATING TO KNOWLEDGE, SKILLS AND OTHER COMPETENCES

1. Student has basic programming skills.

\

SUBJECT OBJECTIVES

C1 Searching, extracting, storing and computer-aided analysis of big data. Understanding its impact and relevance in today's society.

SUBJECT EDUCATIONAL EFFECTS

relating to knowledge:
PEK_W12 knows how to use programming languages and their scientific modules for big data analysis

relating to skills:
PEK_U12 can perform an analysis of big data by making use of a computer

relating to social competences:
PEK_K06 can, without assistance, search for necessary information in the literature, also in foreign languages

PEK_K02 can accurately formulate questions for deeper understanding of a given topic

PROGRAMME CONTENT		
:ecture		Number of hours
Lec1	Introduction to Big Data	2
Lec2	Big data platforms	2
Lec3	Hadoop ecosystem	4
Lec4	Querying big data with Hive	4
Lec5	Big data and machine learning	4
Lec6	In-memory big data platform - Spark	4
Lec7	Linked Big Data	4
Lec8	Big data visualization	2
Lec9	Project presentations	4
	Total hours	**30**

Form of classes - project		Number of hours
Pr1	Practical Preparation and presentations of projects illustrating methods given in the lectures.	30
	Total hours	**30**

TEACHING TOOLS USED
N1. Lecture – traditional method and presentations N2. Student partial project presentation and final presentation N3. Consultations N4. Student's self work – work related to the project development

EVALUATION OF SUBJECT EDUCATIONAL EFFECTS ACHIEVEMENT

Evaluation (F – forming (during semester), P – concluding (at semester end)	Educational effect number	Way of evaluating educational effect achievement
F1	PEK_W12 PEK_U12	mid-term exams
F2	PEK_U12 PEK_K06 PEK_K02	Oral presentations
C P==0.5*F1+0.5*F2		

PRIMARY AND SECONDARY LITERATURE

PRIMARY LITERATURE:

[1] Flach, Peter, Machine Learning, Cambridge University Press, 2012

[2] Holmes, Alex, Hadoop in practice, Manning Publications, 2013

[3] Provost, Foster, Facett, Tom, Data Science for Business. What you need to know about data mining and data-analytic thinking, O'Reilly, 2013

[4] Loshin, David, Big Data Analytics. From Strategic Planning to Enterprise Integration with Tools, Techniques, NoSQL, and Graph, Morgan Kaufmann, 2013

SECONDARY LITERATURE:

[5] http://hadoop.apache.org/, http://spark.apache.org/, http://storm.apache.org/, http://kafka.apache.org/

[6] deRoos, Dirk, Hadoop for Dummies, For Dummies, 2014

SUBJECT SUPERVISOR (NAME AND SURNAME, E-MAIL ADDRESS)