



## Subject card

Subject name and code	Big Data processing frameworks, PG_00045325						
Field of study	Data Engineering						
Date of commencement of studies	October 2024	Academic year of realisation of subject			2026/2027		
Education level	first-cycle studies	Subject group			Optional subject group Subject group related to scientific research in the field of study		
Mode of study	Full-time studies	Mode of delivery			at the university		
Year of study	3	Language of instruction			Polish		
Semester of study	6	ECTS credits			5.0		
Learning profile	general academic profile	Assessment form			exam		
Conducting unit	Department of Software Engineering -> Faculty of Electronics, Telecommunications and Informatics						
Name and surname of lecturer (lecturers)	Subject supervisor	dr Adam Przybyłek					
	Teachers	dr Adam Przybyłek					
Lesson types and methods of instruction	Lesson type	Lecture	Tutorial	Laboratory	Project	Seminar	SUM
	Number of study hours	15.0	0.0	30.0	15.0	0.0	60
	E-learning hours included: 0.0						
Learning activity and number of study hours	Learning activity	Participation in didactic classes included in study plan	Participation in consultation hours		Self-study	SUM	
	Number of study hours	60	6.0		59.0	125	
Subject objectives	The aim of the course is to introduce students to the foundations of Big Data Systems. The course covers 3 frameworks for easily writing applications which process vast amounts of data in-parallel on large clusters of commodity hardware in a reliable, fault-tolerant manner.						
Learning outcomes	Course outcome		Subject outcome		Method of verification		
Subject contents	1. Introduction to Big Data and cloud computing 2. Apache Hadoop 3. Apache Storm 4. Apache Spark 5. Keras						
Prerequisites and co-requisites	Programming in Java and Python.						
Assessment methods and criteria	Subject passing criteria		Passing threshold		Percentage of the final grade		
	lab exam		40.0%		40.0%		
	exam		40.0%		30.0%		
	project		40.0%		30.0%		
Recommended reading	Basic literature		1. <a href="http://hadoop.apache.org/">http://hadoop.apache.org/</a> 2. <a href="http://storm.apache.org/">http://storm.apache.org/</a> 3. <a href="http://spark.apache.org/">http://spark.apache.org/</a>				
	Supplementary literature		1. Hwang, K., Dongarra, J., Fox, G.: Distributed and Cloud Computing: From Parallel Processing to the Internet of Things. Morgan Kaufmann, 2011 2. Karau, H., Konwinski, A., Wendell, P., Zaharia, M.: Learning Spark: Lightning-Fast Big Data Analysis. O'Reilly, 2015 3. Erl, T., Puttini, R., Mahmood, Z.: Cloud Computing: Concepts, Technology, and Architecture. Prentice Hall, 2013 4. Miner, D., Shook, A.: MapReduce Design Patterns: Building Effective Algorithms and Analytics for Hadoop and Other Systems. O'Reilly, 2012				
	eResources addresses		Adresy na platformie eNauczanie:				
Example issues/ example questions/ tasks being completed							

Work placement	Not applicable
----------------	----------------

Document generated electronically. Does not require a seal or signature.