



## Subject card

Subject name and code	Big Data analysis, PG_00045382						
Field of study	Data Engineering						
Date of commencement of studies	October 2021	Academic year of realisation of subject			2024/2025		
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	4	Language of instruction			English		
Semester of study	7	ECTS credits			4.0		
Learning profile	general academic profile	Assessment form			assessment		
Conducting unit	Department of Software Engineering -> Faculty of Electronics, Telecommunications and Informatics						
Name and surname of lecturer (lecturers)	Subject supervisor		dr inż. Wojciech Waloszek				
	Teachers		dr inż. Wojciech Waloszek				
Lesson type and method of instruction	Lesson type	Lecture	Tutorial	Laboratory	Project	Seminar	SUM
	Number of study hours	0.0	0.0	30.0	0.0	0.0	30
	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	30		5.0		65.0	100
Subject objectives	The aim of the course is to familiarize students with the methods of storing and analysis of big data. Practical tools for these tasks are presented.						
Learning outcomes	Course outcome		Subject outcome		Method of verification		
	[K6_W07] Knows the methods of information processing, storage, extraction of data stored in various models including: relational, graph and document ones		Student knows distinguishing features of storing and processing data in large-scaled repositories. A student knows how to acquire data from existing semistructural repositories.		[SW1] Assessment of factual knowledge		
	[K6_U06] Independently solves complex engineering tasks using literature, materials and devices, prepares extensive documentation of the developed solution using appropriate description techniques.		Student designs and executes a process of big data analysis with use of appropriate algorithms		[SU4] Assessment of ability to use methods and tools [SU1] Assessment of task fulfilment		
	[K6_U03] analyses problems and creates appropriate models, data structures and algorithms (including heuristic and numerical ones), assesses their computational complexity, estimates errors of the received solutions		Student designs, creates and implements a large-scaled data repository and assess the computational cost of data processing.		[SU1] Assessment of task fulfilment [SU4] Assessment of ability to use methods and tools		
Subject contents	<ol style="list-style-type: none"> <li>1. Big data characteristics</li> <li>2. Open Linked Data</li> <li>3. Acquiring linked data</li> <li>4. Internet robots and semistructural data analysis</li> <li>5. Storing big data</li> <li>6. Data mining algorithms for big data</li> <li>7. Methods and tools for analysing big data</li> </ol>						
Prerequisites and co-requisites	Basic knowledge about Map-Reduce paradigm.						
Assessment methods and criteria	Subject passing criteria		Passing threshold		Percentage of the final grade		
	Written test		50.0%		30.0%		
	Practical exercises		50.0%		70.0%		

Recommended reading	Basic literature	1. Liu B., "Web data mining", Springer, 2011 2. White T., "Hadoop, the definitive guide", O'Reilly, 2012 3. George L., "HBase, the definitive guide", O'Reilly, 2011
	Supplementary literature	none
	eResources addresses	
Example issues/ example questions/ tasks being completed		
Work placement	Not applicable	