

Subject card

Subject name and code	Implementing Big Data Solutions, PG_00177470						
Field of study	Informatics and Econometrics						
Date of commencement of studies	October 2025		Academic year of realisation of subject		2026/2027		
Education level	Master's studies		Subject group		Obligatory subject group in the field of study 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	2		Language of instruction		English		
Semester of study	4		ECTS credits		5.0		
Learning profile	academic		Assessment form		exam		
Conducting unit	Department of Business Informatics -> Faculty of Management -> Rector						
Name and surname of lecturer (lecturers)	Subject supervisor		dr Patrycja Krauze-Maślankowska				
	Teachers						
Lesson types	Lesson type	Lecture	Tutorial	Laboratory	Project	Seminar	SUM
	Number of study hours	15.0	0.0	45.0	0.0	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		4.0		61.0	125
Subject objectives	Familiarizing students with a comprehensive approach to the acquisition and processing of large data sets.						
	Preparing students to create Big Data solutions.						
Learning outcomes	Course outcome		Subject outcome		Method of verification		
	[liEMU2_U12] The student can adapt, design, create, and operate IT systems that support business entities.		use advanced types of large databases		[SU2] presentation/project/paper/report		
	[liEMU2_U03] The student is able to obtain and verify data from properly selected sources and to collect, process, and visualize it using modern econometrics, informatics or statistics tools.		designs web scraping scripts and open data retrieval		[SU2] presentation/project/paper/report		
	[liEMU2_W05] The student possesses advanced knowledge and understanding of informatics, statistics, and econometrics techniques and tools used to acquire, process, or visualise data to aid in decision-making and verify research hypotheses.		understands the purpose of creating and using Big Data systems		[SW4] test/exam - oral or written		

Subject contents	Lecture		
	Introduction to Big Data, types of data, data division, classifications and technologies		
	Web scraping techniques, generic and dedicated web scraping, legal conditions of web scraping		
	NoSQL databases - collections and documents - creating, saving and downloading data		
	Review of Big Data analytical tools, libraries supporting data processing		
	Apache Hadoop Ecosystem		
	Practical application of Data Mining, Text Mining, Web Mining		
	Supervised and unsupervised machine learning		
	Exercises		
	Web scraping methods - using Python to automatically download data from the Internet		
	Machine learning methods - supervised and unsupervised learning, the use of text and numerical sets		
	Text mining methods automatic extraction of valuable information from text files		
	Collecting large data sets - NoSQL databases, saving website content, creating and selecting queries		
	Working with Open Data type data, using API		
	Data processing from various file formats - JSON, CSV and XML		
Big data processing in Apache Hadoop and Apache Spark - PySpark application, MapReduce algorithms: WordCount analysis, HDFS - Hadoop Distributed File System			
Applications dedicated to web scraping			
Big Data implementation case studies			
Prerequisites and co-requisites	Knowledge of the SQL language and the basics of creating databases.		
Assessment methods and criteria	Subject passing criteria	Passing threshold	Percentage of the final grade
	designing a system using Big Data solutions	50.01%	60.0%
	student's active participation in classes - points earned for correct solving of problems	50.01%	10.0%
	test - assessment of knowledge of Big Data solutions.	50.01%	30.0%

Recommended reading	Basic literature	Glass, R., Callahan, S., (2015) The Big Data-Driven Business: How to Use Big Data to Win Customers, Beat Competitors, and Boost Profits, John Wiley & Sons Documentation Apache Hadoop and Spark: http://hadoop.apache.org, http://spark.apache.org; Python: http://python.org, Education materials on pe.ug.edu.pl
	Supplementary literature	Deitel P., Deitel H., Python dla programistów. Big Data i AI. Studia przypadków, Helion, 2020 Mayer-Schonberger, V., Cukier, K., (2013) Big Data: A Revolution That Will Transform How We Live, Work, and Think, Eamon Dolan/Houghton Mifflin Harcourt
	eResources addresses	
Example issues/ example questions/ tasks being completed	Types of Big Data Systems architectureData Types of Big Data	
Work placement	Not applicable	

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