



Subject card

Subject name and code	Advanced techniques of non-relational databases, PG_00048045						
Field of study	Informatics, Biomedical Engineering, Biomedical Engineering, Biomedical Engineering						
Date of commencement of studies	February 2023		Academic year of realisation of subject		2023/2024		
Education level	second-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		blended-learning		
Year of study	1		Language of instruction		English		
Semester of study	2		ECTS credits		5.0		
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. Teresa Zawadzka				
	Teachers		dr inż. Teresa Zawadzka				
Lesson types and methods of instruction	Lesson type	Lecture	Tutorial	Laboratory	Project	Seminar	SUM
	Number of study hours	15.0	0.0	30.0	30.0	0.0	75
	E-learning hours included: 15.0						
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	75		10.0		40.0	125
Subject objectives	The aim of the course is to familiarize students with analytical data processing in a distributed Hadoop environment.						

Learning outcomes	Course outcome	Subject outcome	Method of verification
	[K7_U03] can design, according to required specifications, and make a complex device, facility, system or carry out a process, specific to the field of study, using suitable methods, techniques, tools and materials, following engineering standards and norms, applying technologies specific to the field of study and experience gained in the professional engineering environment	Student analyzes the problem of building a datawarehouse in a distributed environment and designs this warehouse.	[SU2] Assessment of ability to analyse information [SU1] Assessment of task fulfilment
	[K7_W42] Knows and understands, to an increased extent, the principles and trends in the analysis and design of local and distributed IT systems and the basics of computer modeling and computerization of complex cognitive and decision-making processes.	The student knows the problems of analytical data processing in the Hadoop Environment.	[SW1] Assessment of factual knowledge
	[K7_W03] Knows and understands, to an increased extent, the construction and operating principles of components and systems related to the field of study, including theories, methods and complex relationships between them and selected specific issues - appropriate for the curriculum.	The student knows and understands in depth the structure and operation of Hive and Kylin Tools for analytical data processing in the Hadoop environment.	[SW1] Assessment of factual knowledge
	[K7_K02] is ready to provide critical evaluation of received content and to acknowledge the importance of knowledge in solving cognitive and practical problems	The student is able to critically evaluate the proposed solutions and tools.	[SK5] Assessment of ability to solve problems that arise in practice
	[K7_U04] can apply knowledge of programming methods and techniques as well as select and apply appropriate programming methods and tools in computer software development or programming devices or controllers using microprocessors or programmable elements or systems specific to the field of study, making assessment and critical analysis of the prepared software as well as a synthesis and creative interpretation of information presented with it	The student analyzes the problem of building a datawarehouse in a distributed environment, its requirements and its design, and basing on this information implements it in an appropriate tool.	[SU4] Assessment of ability to use methods and tools [SU3] Assessment of ability to use knowledge gained from the subject [SU2] Assessment of ability to analyse information [SU1] Assessment of task fulfilment
Subject contents	1. HIVE as a data warehouse. 2. Kylin as a data warehouse.		
Prerequisites and co-requisites	Knowledge of relational databases. Basics of the Linux system.		
Assessment methods and criteria	Subject passing criteria	Passing threshold	Percentage of the final grade
	Exam Kylin	50.0%	25.0%
	Project	50.0%	30.0%
	Work during workshop	50.0%	20.0%
	Exam Hive	50.0%	25.0%

Recommended reading	Basic literature	<p>1. Apache Hive Cookbook, Hanish Bansal, Saurabh Chauhan, Shrey Mehrotra, Publisher: Packt Publishing, Release Date: April 2016</p> <p>2. Apache Kylin, http://kylin.apache.org/docs/.</p> <p>3. Apache Hive, https://hive.apache.org/</p>
	Supplementary literature	1. Cloudera documentation, https://www.cloudera.com/products/opensource/apache-hadoop/apache-hive.html .
	eResources addresses	Adresy na platformie eNauczanie:
Example issues/ example questions/ tasks being completed	<p>1. Build a data warehouse in the Hive tool</p> <p>2. Set some tables as external.</p> <p>3. Design at least two partitions.</p> <p>4. Design a data warehouse in the Kylin tool.</p>	
Work placement	Not applicable	