

Institute of Cyberphysical Systems			Semester 1. of the curriculum 2023-24-1			
Name of the subject:	Code of the subject:	Credits:	Weekly hours:			
				lec	sem	lab
Databases and Big Data technologies	NKXDB1EMNF	5	full-time	2	0	2
Responsible person for the subject: Dr. FLEINER Rita			Classification: associate professor			
Subject lecturer(s):						
Prerequisites:						
Way of the assessment: mid-term grade						
Course description						
Goal:	In the course, students learn the principles and implementation of relational database management, the process of database design and modern data management methods. During the course, students will gain insights into the world of non-relational database management and Big Data, and will become familiar with the concepts, procedures and tools of NoSQL and Big Data data storage.					
Course description:	Relational data model, relational algebra, RDBMS architecture, logical and physical data model, database design, normal forms. Database management in Oracle environment database instances, memory structures, transactions. Execution planning, optimization, SQL tuning. Index structures, join methods. NoSQL database types and their operation, their relation to Big Data systems. Understanding the use of MongoDB and Cassandra database management systems: basics, architecture, queries. Big data basics and the Hadoop framework. Apache Spark.					

Lecture schedule	
Education week	Topic
1.	T: Introduction. Knowledge assessment. Relational database systems. L: Basic SQL exercises.
2.	T: Data modelling, single-relationship data model. L: Multi-table queries.
3.	T: Normal forms, dependencies, decomposition of relations. L: DDL, constraints.
4.	T: Relational algebra, relational data model. L: DML, views.
5.	T: Data storage, file organisation, indexes. L: Grouping functions (GROUP BY, HAVING statement parts).
6.	T: Query processing, query optimization. L: Transaction handling.
7.	T: Database tuning. Execution plan, access paths, indexes, join types, CBO statistics, selectivity, cost, materialization, pipelining. L: Execution plan analysis.
8.	T: Database tuning. Execution plan, access paths, indexes, join types, CBO statistics, selectivity, cost, materialization, pipelining. L: Execution plan analysis.
9.	T: NoSQL databases. Cassandra: concepts, architecture, queries. L: Cassandra in practice.
10.	T: NoSQL databases. MongoDB: concepts, architecture, queries. L: MongoDB in practice.
11.	T: Basics of Big data. Hadoop framework. L: Spark in practise.
12.	T: Basics of Big data. Apache Spark. L: Spark in practise.

13.	T: Test (theory + practise)
14.	T: Test replacement
Mid-term requirements	
Conditions for obtaining a mid-term grade/signature	Students have to pass at least 51% in both tests (theory and practise).
Assessment schedule	
Education week	Topic
13.	Theory test, Lab test
14.	Theory test replacement, Lab test replacement
Method used to calculate the <i>mid-term grade</i> (to be filled out only for subjects with mid-term grades)	
The mid-term grade is determined by the sum of the points obtained in the tests.	
Type of the replacement	
Type of the replacement of written test/mid-term grade/signature	Both tests can be replaced in the 14th week and at the beginning of the exam period.
Type of the exam (to be filled out only for subjects with exams)	
Calculation of the exam mark (to be filled only for subjects with exams)	
Final grade calculation methods:	
0% - 51%: failed (1) 52% - 65%: satisfactory (2) 66% - 75%: average (3) 76% - 87%: good (4) 88% - 100%: excellent (5)	
References	
Obligatory:	Jeffrey D. Ullman; Jennifer Widom: Adatbázisrendszerek – Alapvetés (2. kiadás), Panem, 2009. Budapest, ISBN: 9635454815 Elmasri, R., Navathe, S. B.: Fundamentals of Database Systems 7th Edition, ISBN: 978-0133970777
Recommended:	Alex Holmes: Hadoop In Practice, 2nd Edition, September 2014, ISBN 978-1-617-29222-4 Dirk deRoos, Paul C. Zikopoulos, Roman B. Melnyk PhD, Bruce Brown, Rafael Coss: Hadoop for Dummies, 2014 John Wiley & Sons, Inc., Hoboken, New Jersey, ISBN 978-1-118-65220-6
Other references:	