Institute of Cyber-physical Systems			Semester 3. of the curriculum 2026-27-1			
Name of the subject:	Code of the subject:	Credits:	Hours per semester:			
				lec	sem	lab
Database- and Big Data technologies*	NKXAB1EMLF	5	part-time	10	0	10
Responsible person for the subject: Rita FLEINER PhD			Classification: associate professor			
Subject lecturer(s): Zoltan Fried Phd						
Prerequisites:	-					
Way of the assessment:	Midterm grade					

Course description	
Goal:	The aim of the lesson is to familiarize students with advanced database management concepts and procedures, and to develop skills for query writing and Big Data analysis.
Course description:	Relational data models, relational algebra, SQL deep dive. Logical and physical data models, relations. RDBMS design, dependencies, constraints, normal forms, normalization. Triggers and constraints in SQL. Database fundamentals, instance definition, memory structures in db. Transactions. Index types, hashes. SQL tuning. NonSQL database management system i.e. Cassandra, MongoDB. Big data analysis with Apache Spark.

Lecture schedule	
Education week	Topic
1.	Introduction to RDBMS. Relational database systems components analysis. Normalization forms (I, II, III) Basic SQL exercises as knowledge assessment.
2.	Database architecture, indexes, CBO statistics and histograms. Advanced SQL queries using joins, computational costs analysis.
3.	Advanced SQL exercises, merging tables, join types and subqueries, access paths and join algorithms analysis.
4.	Query execution plan interpretation, database tuning introduction and practice with study case, access paths analysis and selection, use of hints as cost reduction and optimization methods.
5.	Homework/RDBMS questions and discussion
6.	1st Mid-Term Test
7.	NoSQL databases: concepts, types. Key-value stores. Data architecture comparison with Oracle RDBMS, System selection criteria based on the CAP theorem.
8.	NoSQL databases. Cassandra: Introduction to the system -- installation, architecture analysis, testing queries.
9.	NoSQL databases. MongoDB: Introduction to the system -- installation, architecture analysis, comparison versus other systems. Basic queries and introduction to pipeline.
10.	NoSQL databases practice case using MongoDB and Cassandra based on provided datasets.
11.	Holiday (Rector's Break)
12.	Basics of Big data. Apache Spark installation and setup. Spark in practice.
13.	Big Data - Practice based on case study using multiple datasets, how to store data in different formats, migrating data.
14.	2nd Mid-Term Test.

Mid-term requirements	
Conditions for obtaining a mid-term grade/signature	**There will be two tests during the course:** 1. **The first test is on the 6th week. Topic: Relational databases.**

	Theory and practice (25%) 2. **The second test is on the 14th week. Topic: Non-SQL DB (*25%*)** **Weekly summary/homework of each lecture (*25%*) *** **Student must solve a homework project in the topics of Cassandra, MongoDB and Spark (25%)** **Prerequisite for obtaining a mid-trem grade: students must pass both tests and all homework with at least 51% each.** ***A student who has missed more than 30% of the classes will not receive a mid-year grade. Late summit for homework project indicates 5%/day minus from result of the homework project.*** **(*) *Student's participation during lectures is considered as part of the Homework.***

Assessment schedule

Education week	Topic
6.	**RDMBS, Database architecture (concepts and queries)**
14.	**Tuning, Advanced SQL (concepts and queries)**

Method used to calculate the *mid-term grade* (to be filled out only for subjects with mid-term grades)

The final grade is formed from result of homeworks, homework project and tests.

Achieved result	Grade
85% -	excellent (5)
74% -	good (4)
63% -	satisfactory (3)
51% -	pass (2)
0% -	failed (1)

Type of the replacement

Type of the replacement of written test/mid-term grade/signature	**If the mid-term exam result is less than 51%, the student may take a make-up exam in week 14. Mid-term grade make-up: once, during the first week of 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)

-

References

Obligatory:	Lecture notes
Recommended:	– 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 – 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:	-