Institute of Cyber-physical Systems			2025/26/2			
Name of the subject:	Code of the subject:	Credits:	Weekly hours:			
				lec	sem	lab
Cloud-based IoT and Big Data platforms	NKXCB1EMNF	4	full-time	2	0	2
Responsible person for the subject: Róbert Lovas Ph.D. habil.			Classification: associate professor			
Subject lecturer(s): Róbert Lovas Ph.D. habil., Márk Emődi						
Prerequisites:						
Way of the assessment:	Midterm grade					

Course description	
Goal:	Within the course framework, students become familiar with major Big Data and Internet of Things (IoT) solutions. The course presents distributed and parallel architectures, their operational mechanisms, the applied technologies, and cloud-based services across various computing platforms, with the aim of efficiently supporting Big Data and IoT application domains, as well as the integration of artificial intelligence (AI) and machine learning into such environments.
Course description:	In the first part of the course, the evolution and key characteristics of Big Data solutions are discussed, including Hadoop- and Spark-based Big Data systems, as well as in-memory and NoSQL database technologies. The course also addresses the theoretical and practical background of management and orchestration solutions for cloud-based Big Data application domains. In addition to introducing IoT and related frameworks, students become familiar with various application areas of data collection. As an extension of the theoretical background, the course covers Lambda, Kappa, and other architectural approaches and provides an overview of the design considerations and impacts of integrating artificial intelligence and machine learning into Big Data and IoT architectures.

Lecture schedule	
Education week	Topic
1.	Lec: Fundamentals of Big Data and Hadoop Lab: Introduction to the World of IoT and Big Data
2.	Lec: Fundamentals of IoT Lab: Docker Container Technology
3.	Lec: Database Scaling and Fundamentals of NoSQL Lab: The Role of Containerization in IoT Solutions
4.	Lec: Document and Graph Databases Lab: Service Management in Containerized Environments
5.	Lec: Column-Oriented and In-Memory Databases Lab: Managing Distributed Systems Using Docker Swarm
6.	Lec: Cloud-Based IoT Lab: The Role of NoSQL Data Models in Data Management
7.	Lec: Cloud-Based Big Data Lab: The Role and Architectural Principles of Kubernetes in Distributed Systems
8.	Lec: **BREAK – EASTER MONDAY** Lab: Implementation of Scalable Applications with Kubernetes
9.	Lec: IoT and Big Data Processing Lab: Data Collection Based on Publish–Subscribe Communication
10.	Lec: Integration of IoT, Big Data, and AI I.

	Lab: Storage and Collection of Time-Series Data
11.	Lec: Integration of IoT, Big Data, and AI II. Lab: **BREAK – TDK (Student Scientific Conference)**
12.	Lec: Integration of IoT, Big Data, and AI III. Lab: Big Data Processing Pipeline
13.	Lec: Midterm Exam Lab: Semester Project Presentation
14.	Lec: Retake Midterm Exam Lab: Retake Semester Project Presentation

Mid-term requirements

Conditions for obtaining a mid-term grade/signature	Successful completion of the midterm, documentation, and presentation of the midterm assignment.

Assessment schedule

Education week	Topic
13.	Midterm exam, Midterm assignment presentation
14.	Midterm exam/assignment retake

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

The completed project work will modify the final result with a -1/0/+1 grade.

Achieved result	**Grade**
89%-100%	excellent (5)
76%-88<%	good (4)
63%-75<%	satisfactory (3)
51%-62<%	pass (2)
0%-50<%	fail (1)

Type of the replacement

Type of the replacement of the written test/mid-term grade/signature	In week 14, students may retake the midterm exam or resubmit the midterm assignment.

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:

References

Obligatory:	Materials published on the Moodle site of the subject.
Recommended:	– Guy Harrison: Next Generation Databases - NoSQL, NewSQL, and Big Data, Apress, 2015 ISBN 978-1-4842-1330-8 – Martin Kleppmann: Designing Data-Intensive Applications – The Big Ideas Behind Reliable, Scalable, and Maintainable Systems, O'Reilly Media, 2017, ISBN 978-1-4493-7332-0 – Tyler Akidau, Slava Chernyak, Reuven Lax: Streaming Systems – The What, Where, When, and How of Large-Scale Data Processing, O'Reilly Media, 2018, ISBN 978-1-4919-0504-1
Other references:	