# SYLLABUS

**Name:**      *Hadoop Ecosystem (AEIAu-DS>SM2HE25)*
**Name in Polish:**
**Name in English:**      *Hadoop Ecosystem*

## Information on course:

**Course offered by department:**    Faculty of Automatic Control, Electronics and Computer Science
**Course for department:**    Silesian University of Technology

**Default type of course examination report:**

ZAL

**Language:**

English

**Course homepage:**

https://platforma2.polsl.pl/rau2/course/view.php?id=916

**Short description:**

The aim of the course is to familiarize students with the concepts related to the processing and analysis of Big data, Big data processing platforms (including Hadoop, Hive, Pig, Spark) and their architecture, methods of data storage and transformation, computational models used on large data processing platforms and programming solutions for data analytics on these platforms.

**Description:**

ECTS: 2
Total workload: 60 hours (30 contact hours, 30 students' own work hours)
Lecture 15h
Laboratory 15h
Students' own work: preparation for tests, reviewing materials, preparation of reports, preparation for classes, elaboration of measurement results


Lecture:
1. Introduction to Big Data. 5V model
2. Big data usage scenarios
3. Architecture of Big Data systems
4. Programming models in the big data ecosystem
5. Scalable big data calculations
6. Big Data Storage Systems (HDFS)
7. Big data analytics platforms (Hadoop, Hive, Pig, Spark)


Labs:
1) Data models and working with NoSQL databases in the cloud
2) Hadoop/Hive/Spark data analysis
3) Data engineering

**Bibliography:**

J. Aven (2017) Hadoop in 24 Hours, Sams Teach Yourself. Sams Publishing; 1 edition (April 17, 2017)
B. Bengfort, J. Kimm (2016) Data Analytics with Hadoop: An Introduction for Data Scientists 1st Edition, O'Reilly Media; 1 edition (June 18, 2016)

**Learning outcomes:**

Knowledge
Understands the basic concepts of big data analytics (post-lecture test, oral talk) - K2A_W15

Knows various data models and formats for big data analytics (post-lecture test, oral talk) - K2A_W15


Skills
He can administer an advanced data processing system (laboratory report) - K2A_U16

The student is able to write a computer program running on a selected large data processing platform (laboratory report, oral talk) - K2A_U01

The student is able to use advanced programming tools to process large data (laboratory report, oral talk) - K2A_U15

**Assessment methods and assessment criteria:**

Lecture
A series of short quizzes after lectures, an oral exam.
Passing criteria: minimum 60% of correct answers

Lab
Lab attendance and assessment min. 3 for the obligatory exercise done, passing all exercises with a positive grade (min. 3).

The syllabus is valid from academic year 2025/26 and its content cannot be changed during the semester.

## Course credits in various terms:

**<without a specific program>**

| Type of credits | Number | First term | Last term |
|---|---|---|---|
| European Credit Transfer System (ECTS) | 2 | 2025/2026-L | |