Big Data

Krzysztof Świder

Rzeszów University of Technology

kswider@prz.edu.pl

https://kswider.kia.prz.edu.pl

Website of the course:

https://kswider.kia.prz.edu.pl/ers

Contents

Topics

1. Introduction

Increase in the volume and complexity of data. Basic trends in data management. The general concept of big data and "V" characteristics. Sources and examples of "big data".

2. Big Data Architecture

Functional requirements. General scheme and characteristics of selected architecture components.

3. The Compute Infrastructure

Handling big data velocity. Cloud computing and data storage. Hadoop and big data. Data storage in HDFS. The MapReduce concept. The Hadoop ecosystem.

Practical work

Using the Apache Spark platform for data processing and analysis.
Run a simple Spark application.