

APPROVED BY

Director of *Institute of Cybernetics*  
 / *Dmitriy M. Sonkin*

## Data Bases

**Field of Study:** *Big Data Solutions*

**Programme name:** *Software Engineering*

**Level of Study:** *Master Degree Programme*

**Year of admission:** 2020

**Semester, year:** *1, 1*

**ECTS:** 3

**Total Hours:** 108

**Contact Hours:** 64

- **Lectures:** 16
- **Labs:** 32


**Assessment:** *credit test & project, exam*

**Department:** *Software Engineering*

**Head of Department**

  
\_\_\_\_\_ / V.S. Sherstnev

**Instructor(s)**

  
\_\_\_\_\_ / V.V. Sokolova



## Course Name

### Course Overview

<b>Course Objectives</b>	<i>The objective of the course is to gain the ability to employ the latest tools, technologies and techniques required to analyze, debug, iterate and optimize the analysis to infer actionable insights from Big Data.</i>
<b>Learning Outcomes</b>	<i>The learning outcomes are to apply key technologies used in acquiring, organizing, storing, and analyzing big data</i>
<b>Course Outline</b>	<p><i>Lectures are devoted to the methods for acquiring, storing and analysing big data.</i></p> <p><i>In labs students have an opportunity to apply the latest tools to analyze, debug and optime the analysis of big data</i></p> <p><i>Practice works are aimed to make a group project concerning with applying the technologies relevant for analysis of big data in the specific knowledge domain.</i></p>
<b>Prerequisites (if available)</b>	<i>Introduction to Big Data, Data Analysis Methods, Big Data Programming Tools</i>
<b>Course Structure</b>	<ol style="list-style-type: none"> <li><i>1. Discovering data from big data management systems</i></li> <li><i>2. Analyzing the connections between data management operations and the big data processing patterns</i></li> <li><i>3. Vizualizing the results of the analysis</i></li> <li><i>4. Presenting big data of the specific knowledge domain</i></li> </ol>
<b>Facilities and Equipment</b>	<i>Apache Hadoop, Apache Spark Core</i>
<b>Grading Policy</b>	<p><i>In accordance with TPU rating system we use:</i></p> <ul style="list-style-type: none"> <li><i>- Current assessment which is performed on a regular basis during the semester by scoring the quality of mastering of theoretical material and the results of practical activities (performance tests, perform tasks, problem solving). Max score for current assessment is 60 points, min – 40 points.</i></li> <li><i>- Course final assessment (exam/ credit test) is performed at the end of the semester. Max score for course final assessment is 40 points, min – 22 points.</i></li> </ul> <p><i>The final rating is determined by summing the points of the current assessment during the semester and exam (credit test) scores at the end of the semester. Maximum overall rating corresponds to 100 points, min pass score is 32.</i></p>
<b>Course Policy</b>	<i>Students are expected to make and present the results of all tasks provided in practice and lab works.</i>
<b>Teaching Aids and Resources</b>	<p><i>Compulsory Readings:</i></p> <ol style="list-style-type: none"> <li><i>1. Data Science &amp; Big Data Analytics: Discovering, Analyzing, Visualizing and Presenting Data // EMC Education Services, John Wiley &amp; Sons, Inc, 2015, 432 p.</i></li> <li><i>2. Analytics in a Big Data World: The Essential Guide to Data Science and its Applications // Bart Baesens, John Wiley &amp; Sons, 2014, 256 p.</i></li> </ol>



	<p>3. <i>Big-Data Analytics and Cloud Computing: Theory, Algorithms and Applications</i> // Marcello Trovati, Richard Hill, Ashiq Anjum, Shao Ying Zhu, Lu Liu, Springer, 2016, 169 p.</p> <p><i>Additional Readings:</i></p> <p>1. <i>Big Data and Business Analytics</i> // Jay Liebowitz, CRC Press, 2016, 304 p.</p> <p>2. <i>High-Performance Big-Data Analytics: Computing Systems and Approaches</i> // Pethuru Raj, Anupama Raman, Dhivya Nagaraj, Siddhartha Duggirala, Springer, 2015, 428 p.</p>
<b>Instructor (-s)</b>	Veronica Sokolova, veronica@tpu.ru