Syllabus of the course

«*High Performance Systems of Processing and Analysis of Big Data*»

Спеціальність	*F3 Computer sciences*
Освітня програма	*Computer sciences*
Освітній рівень	*second (Master)) level of higher education*
Статус дисципліни	*mandatory*
Мова викладання, навчання та оцінювання	*English*
Курс / семестр	*first year, 1 semester*
Кількість кредитів ЄКТС	*5*
Розподіл годин за формами освітнього процесу та видами навчальних занять	*Lectures – 14 hours.* *Practical studies – 0 hours.* *Laboratory studies – 26 hours.* *Self-study – 110 hours.*
Форма семестрового контролю	*Exam*
Кафедра	*Department of Information Systems* *phone: (057) 702 18 31 (add. 3-16)* *website: https://kafis.hneu.net/*
Викладач	*Minukhin Serhii Volodymyrovych, doctor of Technical Sciences, Professor*
Контактна інформація викладача	*Minukhin S.V. serhii.minukhin@@hneu.net*
Дні навчальних занять	*Lectures: according to the current class schedule* *Laboratory studies: according to the current class schedule*
Консультації	*At the Department of Information Systems, offline, according to the schedule, individual, PNS chat.*

The purpose of teaching the is to provide higher education students with a system of theoretical knowledge and acquire practical skills to understand the essence of problems that arise when using big data, modern approaches and tools for their processing and analysis.

Learning objectives and skill

Prerequisites	Postrequsites
	The course work: Development of computer information systems
	Complex training
	Diploma work

Structural and logical scheme of the course

Content module 1. *Basic concepts, essence and features of big data. Principles of organizing the construction of systems for working with big data*

Topic 1. Concepts, characteristics of big data and their processing systems.

Topic 2. Modern big data processing systems. Composition of components and their purpose.

Topic 3. Apache Hadoop: a framework for processing big data. Basic components for building Hadoop: Google's MapReduce, Google File System.

Topic 4. Architecture of Apache Hadoop.

Content Module 2. *Apache Spark: A Universal Platform for Big Data Processing and Analytics*

Topic 5. Architecture of Apache Spark.

Topic 6. Apache Spark deployment modes.

Topic 7. Scheduling tasks in Apache Spark.
Topic 8. Working with databases and data stores in SparkSQL. RDD, Dataframe and Dataset.
Topic 9. Deployment and configuration of Apache Spark and Apache Hadoop frameworks in distributed and virtual environments.

Teaching environment (software)

Distance learning tools: Personalized learning systems website: https://pns.hneu.edu.ua Library: http://library.hneu.edu.ua Repository: http://www.repository.hneu.edu.ua University classrooms (Kharkiv, 9A Nauky Ave.)

Multimedia equipment: projector, laptop / computer, Internet access, software: Microsoft Windows, Microsoft Office, Vagrant, Virtualbox, OS Ubuntu, Apache Spark

Assessment system

An university uses 100 ball story system of evaluation of results of studies of bread-winners of higher education. Current control comes true during realization of lecture and laboratory employments and has for an object verification of level of preparedness of bread-winner of higher education to implementation of concrete work and estimated by the sum of the c ollected points. Final control includes semester control that is conducted in a form to examination. Maximally possible amount of points for current control during a semester for discipline form of c ontrol of that examination - 60 and minimum possible amount of points – 35.

More detailed information on assessment and grading system is given in the technological card of the course.

Course policies

Teaching of the academic discipline is based on the principles of academic integrity.

Violation of academic integrity includes academic plagiarism, fabrication, falsification, cheating, deception, bribery, and biased assessment.

Education seekers may be brought to the following academic responsibility for breach of academic integrity: repeated assessment of the corresponding type of learning activity.

More detailed information about competencies, learning outcomes, teaching methods, assessment forms, self-study is given in the Course program