

Документ подписан Министерством науки и высшего образования Российской Федерации
Информация о владельце:
ФИО: Макаренко Елена Николаевна
Должность: Ректор
Дата подписания: 20.05.2024 11:04:49
Уникальный программный ключ:
c098bc0c1041cb2a4cf926cf171d6715d99a6ae00adc8e27b55cbe1e2dbd7c78

УТВЕРЖДАЮ
Начальник отдела лицензирования и
аккредитации
_____ Чаленко К.Н.
« ____ » _____ 20__ г.

**Рабочая программа дисциплины
Системы больших данных**

основная профессиональная образовательная программа по направлению 02.03.02
Фундаментальная информатика и информационные технологии
02.03.02.01 "Теоретические основы информатики и компьютерные науки"

Для набора 2021 года

Квалификация
Бакалавр

КАФЕДРА **Информационных систем и прикладной информатики****Распределение часов дисциплины по семестрам**

Семестр (<Курс>.<Семестр на курсе>)	7 (4.1)		Итого	
	16			
Неделя	16			
Вид занятий	УП	РП	УП	РП
Лекции	6	6	6	6
Лабораторные	6	6	6	6
Итого ауд.	12	12	12	12
Контактная работа	12	12	12	12
Сам. работа	92	92	92	92
Часы на контроль	4	4	4	4
Итого	108	108	108	108

ОСНОВАНИЕ

Учебный план утвержден учёным советом вуза от 31.08.2021 протокол № 1.

Программу составил(и): д.э.н., доц., Щербаков С.М. _____

Зав. кафедрой: д.э.н., доцент Щербаков С.М. _____

Методическим советом направления: д.э.н., проф., Тищенко Е.Н. _____

1. ЦЕЛИ ОСВОЕНИЯ ДИСЦИПЛИНЫ

1.1	формирование знаний и навыков работы по принципам, методам, технологиям и инструментам использования больших данных в информационных системах в экономике.
-----	--

2. ТРЕБОВАНИЯ К РЕЗУЛЬТАТАМ ОСВОЕНИЯ ДИСЦИПЛИНЫ

ПК-3: Способность собирать, обрабатывать и интерпретировать экспериментальные данные, необходимые для проектной и производственно-технологической деятельности; способность к разработке новых алгоритмических, методических и технологических решений в конкретной сфере профессиональной деятельности

В результате освоения дисциплины обучающийся должен:

Знать:
технологии хранения, обработки и анализа больших данных
Уметь:
анализировать и выбирать методы построения информационных систем больших данных
Владеть:
навыками разработки информационных систем больших данных на основе нереляционных баз данных и распределенных систем хранения

3. СТРУКТУРА И СОДЕРЖАНИЕ ДИСЦИПЛИНЫ

Код занятия	Наименование разделов и тем /вид занятия/	Семестр / Курс	Часов	Компетенции	Литература
	Раздел 1. «Основы построения и использования систем больших данных»				
1.1	Тема 1.1 «Основы систем больших данных» Понятие больших данных. Особенности сбора, хранения, обработки и анализа больших массивов данных. Источники больших данных. Использование больших данных в науке, бизнесе, государственном управлении. /Лек/	7	2	ПК-3	Л1.1 Л1.2Л2.1 Л2.2 Л2.3
1.2	Тема 1.1 «Основы систем больших данных» Средства построения распределенных информационных систем для BigData. Обзор возможностей. Выбор инструментария для разработки. Выполнение лабораторных заданий с использованием Eclipse, LibreOffice. /Лаб/	7	2	ПК-3	Л1.1 Л1.2Л2.1 Л2.2 Л2.3
1.3	Тема: Работа с распределенными информационными системами Распределенность. Клиент-серверная архитектура. Распределенная база данных. Коммуникация. Масштабируемость. Отказоустойчивость. Безопасность. Управление ресурсами. Согласованность данных. Мониторинг и управление. /Ср/	7	46	ПК-3	Л1.1 Л1.2Л2.1 Л2.2 Л2.3
	Раздел 2. «Разработка и использование приложений на основе распределенных баз данных»				
2.1	Тема 2.1 «Базы данных NoSQL» Варианты построения распределенных баз данных, репликация, фрагментация. Согласованность. CAP-теорема. Классы NoSQL баз данных. Примеры СУБД NoSQL. Семейства столбцов. Графовые СУБД. /Лек/	7	4	ПК-3	Л1.1 Л1.2Л2.1 Л2.2 Л2.3
2.2	Тема 2.1 «Базы данных NoSQL» Изучение основных классов NoSQL СУБД, графовых, мультиколоночных, документо-ориентированных, типа "имя=значение". Проектирование и разработка графовой базы данных. Поисковые запросы. Выполнение лабораторных заданий с использованием Eclipse, LibreOffice. /Лаб/	7	4	ПК-3	Л1.1 Л1.2Л2.1 Л2.2 Л2.3
2.3	Тема: Документно-ориентированные распределенные СУБД Прозрачность расположения. Автономность узлов. Обработка распределенных запросов. Выполнение распределенных транзакций. /Ср/	7	46	ПК-3	Л1.1 Л1.2Л2.1 Л2.2 Л2.3

2.4	/Зачёт/	7	4	ПК-3	Л1.1 Л1.2Л2.1 Л2.2 Л2.3
-----	---------	---	---	------	----------------------------

4. ФОНД ОЦЕНОЧНЫХ СРЕДСТВ

Структура и содержание фонда оценочных средств для проведения текущей и промежуточной аттестации представлены в Приложении 1 к рабочей программе дисциплины.

5. УЧЕБНО-МЕТОДИЧЕСКОЕ И ИНФОРМАЦИОННОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ

5.1. Основная литература

	Авторы, составители	Заглавие	Издательство, год	Колич-во
Л1.1	Слюсаренко П. И.	Распределенные СУБД: практическое пособие	Москва: Лаборатория книги, 2012	https://biblioclub.ru/index.php?page=book&id=142013 неограниченный доступ для зарегистрированных пользователей
Л1.2	Железнов, М. М.	Методы и технологии обработки больших данных: учебно-методическое пособие	Москва: МИСИ-МГСУ, ЭБС АСВ, 2020	https://www.iprbookshop.ru/101802.html неограниченный доступ для зарегистрированных пользователей

5.2. Дополнительная литература

	Авторы, составители	Заглавие	Издательство, год	Колич-во
Л2.1	Лихачева Г. Н., Гаспарян М. С.	Информационные системы и технологии: учебно-методический комплекс	Москва: Евразийский открытый институт, 2011	https://biblioclub.ru/index.php?page=book&id=90543 неограниченный доступ для зарегистрированных пользователей
Л2.2	Емельянов А. А.	Прикладная информатика: журнал	Москва: Синергия ПРЕСС, 2006	https://biblioclub.ru/index.php?page=book&id=120298 неограниченный доступ для зарегистрированных пользователей
Л2.3	Петрухнова, Г. В.	Введение в распределенные системы: учебное пособие	Воронеж: Воронежский государственный технический университет, ЭБС АСВ, 2021	https://www.iprbookshop.ru/111462.html неограниченный доступ для зарегистрированных пользователей

5.3 Профессиональные базы данных и информационные справочные системы

ИСС «КонсультантПлюс»

ИСС «Гарант» <http://www.internet.garant.ru/>

Национальная электронная библиотека (НЭБ), <https://rusneb.ru/>

5.4. Перечень программного обеспечения

Eclipse

LibreOffice

5.5. Учебно-методические материалы для студентов с ограниченными возможностями здоровья

При необходимости по заявлению обучающегося с ограниченными возможностями здоровья учебно-методические материалы предоставляются в формах, адаптированных к ограничениям здоровья и восприятия информации. Для лиц с нарушениями зрения: в форме аудиофайла; в печатной форме увеличенным шрифтом. Для лиц с нарушениями слуха: в форме электронного документа; в печатной форме. Для лиц с нарушениями опорно-двигательного аппарата: в форме электронного документа; в печатной форме.

6. МАТЕРИАЛЬНО-ТЕХНИЧЕСКОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ (МОДУЛЯ)

Помещения для проведения всех видов работ, предусмотренных учебным планом, укомплектованы необходимой специализированной учебной мебелью и техническими средствами обучения. Для проведения лекционных занятий используется демонстрационное оборудование. Лабораторные занятия проводятся в компьютерных классах, рабочие места в которых оборудованы необходимыми лицензионными программными средствами и выходом в Интернет.

7. МЕТОДИЧЕСКИЕ УКАЗАНИЯ ДЛЯ ОБУЧАЮЩИХСЯ ПО ОСВОЕНИЮ ДИСЦИПЛИНЫ (МОДУЛЯ)

Методические указания по освоению дисциплины представлены в Приложении 2 к рабочей программе дисциплины.

ФОНД ОЦЕНОЧНЫХ СРЕДСТВ

1. Описание показателей и критериев оценивания компетенций на различных этапах их формирования, описание шкал оценивания

1.1. Показатели и критерии оценивания компетенций:

ЗУН, составляющие компетенцию	Показатели оценивания	Критерии оценивания	Средства оценивания
ПК-3: Способность собирать, обрабатывать и интерпретировать экспериментальные данные, необходимые для проектной и производственно-технологической деятельности; способность к разработке новых алгоритмических, методических и технологических решений в конкретной сфере профессиональной деятельности			
З. технологии хранения, обработки и анализа больших данных	знает основные понятия и определения, методы, алгоритмы и технологии	полнота и содержательность ответа умение приводить примеры	Т – тест (1-10), З – вопросы к зачету (1-12)
У. анализировать и выбирать методы построения информационных систем больших данных	выполняет лабораторные задания, отвечает на вопросы, умеет применять полученные знания на практике	полнота и содержательность ответа умение приводить примеры умение самостоятельно находить решение поставленных задач	ЛЗ – лабораторные задания (1-2)
В. навыками разработки информационных систем больших данных на основе нереляционных баз данных и распределенных систем хранения	проводит обобщенный анализ информации и обработку данных	полнота и содержательность ответа умение приводить примеры умение самостоятельно находить решение поставленных задач	ЛЗ – лабораторные задания (1-2)

1.2 Шкалы оценивания:

Текущий контроль успеваемости и промежуточная аттестация осуществляется в рамках накопительной балльно-рейтинговой системы в 100-балльной шкале:

50-100 баллов (зачет);

0-49 баллов (незачет).

2. Типовые контрольные задания или иные материалы, необходимые для оценки знаний, умений, навыков и (или) опыта деятельности, характеризующих этапы формирования компетенций в процессе освоения образовательной программы

Вопросы к зачету

- 1) Понятие больших данных
- 2) Особенности сбора, хранения, обработки и анализа больших данных
- 3) Требования к распределенным информационным системам
- 4) Средства построения распределенных информационных систем
- 5) Технология Map-Reduce
- 6) Система Apache Hadoop
- 7) Базы данных NoSQL. Особенности, классификация
- 8) Возможности NoSQL-баз данных по обеспечению целостности, доступности скорости обработки информации. CAP-теорема
- 9) Способы репликации и кластеризации баз данных
- 10) Документно-ориентированные базы данных
- 11) Возможности СУБД MongoDB
- 12) Работа с документно-ориентированными БД на языке JSON

Зачетное задание включает два вопроса – один теоретический вопрос и одно практико-ориентированное задание из числа приведенных ниже лабораторных заданий.

Критерии оценивания:

- 50-100 баллов («зачет») – изложенный материал фактически верен, наличие глубоких исчерпывающих знаний в объеме пройденной программы дисциплины в соответствии с поставленными программой курса целями и задачами обучения; правильные, уверенные действия по применению полученных знаний на практике, грамотное и логически стройное изложение материала при ответе, усвоение основной и знакомство с дополнительной литературой; наличие твердых и достаточно полных знаний в объеме пройденной программы дисциплины в соответствии с целями обучения, правильные действия по применению знаний на практике, четкое изложение материала, допускаются отдельные логические и стилистические погрешности, обучающийся усвоил основную литературу, рекомендованную в рабочей программе дисциплины; наличие твердых знаний в объеме пройденного курса в соответствии с целями обучения, изложение ответов с отдельными ошибками, уверенно исправленными после дополнительных вопросов; правильные в целом действия по применению знаний на практике;

- 0-49 баллов («незачет») – ответы не связаны с вопросами, наличие грубых ошибок в ответе, непонимание сущности излагаемого вопроса, неумение применять знания на практике, неуверенность и неточность ответов на дополнительные и наводящие вопросы.

Тест

- 1 Основной признак / признаки больших данных (Big Data), один ответ
 - a) большой объем (volume)
 - b) высокая скорость изменения (velocity)
 - c) многообразие (variety)
 - d) все три характеристики volume, velocity, variety

- 2 Какие базы данных НЕ относятся к NoSQL
 - a) графовые
 - b) документо-ориентированные
 - c) вида ключ-значение
 - d) реляционные

- 3 Основные источники Больших данных
 - a) геопозиционирование мобильных устройств
 - b) log-файлы web-серверов
 - c) социальные сети
 - d) все перечисленное

- 4 Какая СУБД относится к NoSQL?
 - a) MS SQL
 - b) MySQL
 - c) Oracle
 - d) Mongo DB

- 5 Как называется программный фреймворк, предназначенный для распределенных систем и предполагающий отражение данных и их последующую свертку?
 - a) XML
 - b) JSON
 - c) SQL
 - d) MapReduce

- 6 Могут ли документы одной коллекции документо-ориентированной СУБД (например, Mongo DB) отличаться по составу полей?
 - a) Да
 - b) Нет
 - c) Только если это определено схемой данных Mongo
 - d) Только если располагаются на разных узлах кластера

7 Что из перечисленного является платформой построения приложений для обработки больших объемов данных?

- a) Hadoop
- b) MS Excel
- c) JSON
- d) JavaScript

8 Что такое фрагментация (sharding) в СУБД Mongo DB?

- a) Синоним репликации
- b) Способ резервного копирования данных
- c) Разделение базы данных для ее расположения на нескольких узлах
- d) Периодический процесс удаления устаревших фрагментов БД

9 Что такое репликация в распределенных системах хранения данных?

- a) Размещение копий данных на различных узлах для ускорения доступа на чтение и обеспечение отказоустойчивости и сохранения данных
- b) Способ резервного копирования данных
- c) Разделение базы данных для ее расположения на нескольких узлах, синоним фрагментации
- d) Периодический процесс удаления устаревших фрагментов БД

10 В экосистеме языка Python для обработки массивов данных, их очистки и преобразования используется библиотека

- a) Pandas
- b) Random
- c) Matplotlib
- d) PyTorch

Критерии оценивания:

Для одного обучающегося формируется вариант, содержащий 10 вопросов.

Правильный ответ на один вопрос – 1 балл, неправильный – 0 баллов.

Максимальное количество баллов за тест – 10.

Лабораторные задания

Лабораторное задание 1

Тема 1.1 «Основы систем больших данных»

Средства построения распределенных информационных систем для BigData. Обзор возможностей. Выбор инструментария для разработки.

Лабораторное задание 2

Тема 2.1 «Базы данных NoSQL»

Изучение основных классов NoSQL СУБД, графовых, мультиконочных, документо-ориентированных, типа "имя=значение". Проектирование и разработка графовой базы данных. Поискные запросы.

Критерии оценивания (для каждого задания):

34-45 б. – задание выполнено верно;

22-33 б. – при выполнении задания были допущены неточности, не влияющие на результат;

12-21 б. – при выполнении задания были допущены ошибки;

0-11 б. – при выполнении задания были допущены существенные ошибки.

Максимальное количество баллов за лабораторные задания – 90 (2 задания по 45 баллов).

3. Методические материалы, определяющие процедуры оценивания знаний, умений, навыков и (или) опыта деятельности, характеризующих этапы формирования компетенций

Процедуры оценивания включают в себя текущий контроль и промежуточную аттестацию.

Текущий контроль успеваемости проводится с использованием оценочных средств, представленных в п. 2 данного приложения. Результаты текущего контроля доводятся до сведения студентов до промежуточной аттестации.

Промежуточная аттестация проводится в форме зачета.

Зачет проводится по расписанию промежуточной аттестации. Количество вопросов в зачетном задании – 2 (один теоретический вопрос и одно практико-ориентированное задание). Объявление результатов производится в день зачета. Результаты аттестации заносятся в зачетную ведомость и зачетную книжку студента. Студенты, не прошедшие промежуточную аттестацию по графику сессии, должны ликвидировать задолженность в установленном порядке.

МЕТОДИЧЕСКИЕ УКАЗАНИЯ ПО ОСВОЕНИЮ ДИСЦИПЛИНЫ

Учебным планом предусмотрены следующие виды занятий:

- лекции;
- лабораторные занятия;

В ходе лекционных занятий рассматриваются основные теоретические вопросы, даются рекомендации для самостоятельной работы и подготовке к лабораторным занятиям и практическим занятиям.

В ходе лабораторных углубляются и закрепляются знания студентов по ряду рассмотренных на лекциях вопросов, развиваются навыки практической работы.

При подготовке к лабораторным занятиям каждый студент должен:

- изучить рекомендованную учебную литературу;
- изучить конспекты лекций;
- подготовить ответы на все вопросы по изучаемой теме.

В процессе подготовки к лабораторным студент может воспользоваться консультациями преподавателя.

Вопросы, не рассмотренные на лекциях, лабораторных занятиях, должны быть изучены студентами в ходе самостоятельной работы. Контроль самостоятельной работы студентов над учебной программой курса осуществляется в ходе занятий методом теста и выполнения лабораторных заданий. В ходе самостоятельной работы каждый студент обязан прочитать основную и по возможности дополнительную литературу по изучаемой теме, дополнить конспекты лекций недостающим материалом, выписками из рекомендованных первоисточников, выделить непонятные термины, найти их значение в энциклопедических словарях.

Студент должен готовиться к предстоящему лабораторному занятию по всем обозначенным в рабочей программе дисциплины вопросам.

Для подготовки к занятиям, текущему контролю и промежуточной аттестации студенты могут воспользоваться электронно-библиотечными системами. Также обучающиеся могут взять на дом необходимую литературу на абонементе университетской библиотеки или воспользоваться читальными залами.