

Документ подписан простой электронной подписью  
Информация о владельце:  
ФИО: Макаренко Елена Николаевна  
Должность: Ректор  
Дата подписания: 2024.02.28 10:01  
Уникальный программный ключ:  
c098bc0c1041cb2a4cf926cf171d6715d99a6ae00adc8e27b55cbe1e2dbd7c78

Министерство науки и высшего образования Российской Федерации  
Федеральное государственное бюджетное образовательное учреждение высшего образования «Ростовский государственный экономический университет (РИНХ)»

УТВЕРЖДАЮ  
Директор Института магистратуры  
Иванова Е.А.  
«30» 08 2024 г.

**Рабочая программа дисциплины  
Технологии Big Data**

Направление 09.04.03 Прикладная информатика  
магистерская программа 09.04.03.01 "Информационные системы и технологии в бизнесе"

Для набора 2021 года

Квалификация  
магистр

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

Курс	2		Итого	
	уп	рп		
Лекции	6	6	6	6
Лабораторные	10	10	10	10
Практические	4	4	4	4
Итого ауд.	20	20	20	20
Контактная работа	20	20	20	20
Сам. работа	295	295	295	295
Часы на контроль	9	9	9	9
Итого	324	324	324	324

**ОСНОВАНИЕ**

Учебный план утвержден учёным советом вуза от 30.08.2021 протокол № 1.

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

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

Методическим советом направления: д.э.н., доц., Щербаков С.М. *сш*

### 1. ЦЕЛИ ОСВОЕНИЯ ДИСЦИПЛИНЫ

1.1	освоить принципы, методы, технологии и инструменты использования больших данных в информационных системах в экономике.
-----	--

### 2. ТРЕБОВАНИЯ К РЕЗУЛЬТАТАМ ОСВОЕНИЯ ДИСЦИПЛИНЫ

**ПК-10:**Способен осуществлять экспертную поддержку разработки прототипов ИС

**ПК-3:**Способен проводить научные эксперименты, оценивать результаты исследований

#### В результате освоения дисциплины обучающийся должен:

<b>Знать:</b>
особенности работы с большими неструктурированными и слабоструктурированными данными (соотнесено с индикатором ПК-3.1) принципы технологии NoSQL (соотнесено с индикатором ПК-10.1)
<b>Уметь:</b>
настраивать и организовывать NoSQL базы данных (соотнесено с индикатором ПК-3.2) выбирать NoSql СУБД для решения прикладной задачи (соотнесено с индикатором ПК-10.2)
<b>Владеть:</b>
технологиями и языками манипулирования данными (соотнесено с индикатором ПК-3.3) графовыми и документо-ориентированными СУБД (соотнесено с индикатором ПК-10.3)

### 3. СТРУКТУРА И СОДЕРЖАНИЕ ДИСЦИПЛИНЫ

Код занятия	Наименование разделов и тем /вид занятия/	Семестр / Курс	Часов	Компетенции	Литература
	<b>Раздел 1. Основы построения и использования систем больших данных</b>				
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/	2	2	ПК-3 ПК-10	Л1.1 Л1.2Л2.1 Л2.2 Л2.3
1.2	Тема 1.2 "Методы работы с распределенными информационными системами" Использование фреймворка Map-Reduce в распределенной среде. Реализации Map-Reduce. Состав и возможности программного комплекса Apache Hadoop. Языки поисковых запросов для Hadoop. /Лек/	2	2	ПК-3 ПК-10	Л1.1 Л1.2Л2.1 Л2.2 Л2.3
1.3	Тема 1.1 "Основы систем больших данных" Средства построения распределенных информационных систем для BigData. Обзор возможностей. /Лаб/	2	2	ПК-3 ПК-10	Л1.1 Л1.2Л2.1 Л2.2 Л2.3
1.4	Тема 1.2 "Методы работы с распределенными информационными системами" Изучение и конфигурирование программного комплекса Apache Hadoop. Размещение набора данных по заданной тематике. Построение поисковых запросов на языке Pig /Лаб/	2	2	ПК-3 ПК-10	Л1.1 Л1.2Л2.1 Л2.2 Л2.3
1.5	Методы работы с распределенными информационными системами /Ср/	2	98	ПК-3 ПК-10	Л1.1 Л1.2Л2.1 Л2.2 Л2.3
1.6	Базы данных NoSQL /Ср/	2	98	ПК-3 ПК-10	Л1.1 Л1.2Л2.1 Л2.2 Л2.3
	<b>Раздел 2. Разработка и использование приложений на основе распределенных баз данных</b>				
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/	2	2	ПК-3 ПК-10	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 в СУБД Neo4j на заданную тему. Поисковые запросы на языке Cypher. Среда разработки Eclipse. /Лаб/	2	4	ПК-3 ПК-10	Л1.1 Л1.2Л2.1 Л2.2 Л2.3
2.3	Тема 2.2 "Документно-ориентированные распределенные СУБД" Работа с MongoDB. Запросы на выборку и модификацию. Использование драйверов. Настройка фрагментации. Использование Map-Reduce /Лаб/	2	2	ПК-3 ПК-10	Л1.1 Л1.2Л2.1 Л2.2 Л2.3
2.4	Тема 2.1 "Базы данных NoSQL" Проектирование и построение графовой базы данных в СУБД Neo4j /Пр/	2	2	ПК-3 ПК-10	Л1.1 Л1.2Л2.1 Л2.2 Л2.3
2.5	Тема 2.2 "Документно-ориентированные распределенные СУБД" Проектирование структуры документо-ориентированной базы данных. /Пр/	2	2	ПК-3 ПК-10	Л1.1 Л1.2Л2.1 Л2.2 Л2.3
2.6	Документно-ориентированные распределенные СУБД /Ср/	2	99	ПК-3 ПК-10	Л1.1 Л1.2Л2.1 Л2.2 Л2.3
2.7	/Экзамен/	2	9	ПК-3 ПК-10	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	<a href="https://biblioclub.ru/index.php?page=book&amp;id=142013">https://biblioclub.ru/index.php?page=book&amp;id=142013</a> неограниченный доступ для зарегистрированных пользователей
Л1.2	Пол Дейтел, Харви Дейтел	Python: Искусственный интеллект, большие данные и облачные вычисления	Санкт-Петербург: Питер, 2021	<a href="https://ibooks.ru/reading.php?short=1&amp;productid=371701">https://ibooks.ru/reading.php?short=1&amp;productid=371701</a> неограниченный доступ для зарегистрированных пользователей

##### 5.2. Дополнительная литература

	Авторы, составители	Заглавие	Издательство, год	Колич-во
Л2.1	Емельянов А. А.	Прикладная информатика: журнал	Москва: Синергия ПРЕСС, 2010	<a href="https://biblioclub.ru/index.php?page=book&amp;id=120321">https://biblioclub.ru/index.php?page=book&amp;id=120321</a> неограниченный доступ для зарегистрированных пользователей
Л2.2		Распределенные базы данных: учебное пособие	Ставрополь: Северо-Кавказский Федеральный университет (СКФУ), 2015	<a href="https://biblioclub.ru/index.php?page=book&amp;id=457594">https://biblioclub.ru/index.php?page=book&amp;id=457594</a> неограниченный доступ для зарегистрированных пользователей

	Авторы, составители	Заглавие	Издательство, год	Колич-во
Л2.3	Братченко, Н. Ю.	Распределенные базы данных: лабораторный практикум	Ставрополь: Северо-Кавказский федеральный университет, 2014	<a href="http://www.iprbookshop.ru/63129.html">http://www.iprbookshop.ru/63129.html</a> неограниченный доступ для зарегистрированных пользователей

### 5.3 Профессиональные базы данных и информационные справочные системы

Консультант+

Гарант

Центральная база статистических данных <https://www.gks.ru/dbscripts/cbsd/dbinet.cgi>

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

### 5.4. Перечень программного обеспечения

Eclipse

Hadoop

MongoDB

### 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: Способен проводить научные эксперименты, оценивать результаты исследований			
З. особенности работы с большими неструктурированными и слабоструктурированными данными	Основные признаки BigData (3V), инструментарий обработки и анализа BigData	полнота и содержательность ответа умение приводить примеры	О – опрос (вариант 1), З – вопросы к зачету (1-3), Э – вопросы к экзамену (1-6)
У. настраивать и организовывать NoSQL базы данных	Хранение и обработка больших объемов данных в различных типах NoSQL СУБД	полнота и содержательность ответа умение приводить примеры умение самостоятельно находить решение поставленных задач	ЛЗ – лабораторные задания (1-2), ПЗ – практические задания (1)
В. технологиями и языками манипулирования данными	Фреймворк map-reduce, языки доступа к данным	полнота и содержательность ответа умение приводить примеры умение самостоятельно находить решение поставленных задач	ЛЗ – лабораторные задания (1-2), ПЗ – практические задания (1)
ПК-10: Способен осуществлять экспертную поддержку разработки прототипов ИС			
З. принципы технологии NoSQL	Виды NoSQL баз данных	полнота и содержательность ответа умение приводить примеры	О – опрос (вариант 1), З – вопросы к зачету (4-6), Э – вопросы к экзамену (7-12)
У. выбирать NoSql СУБД для решения прикладной задачи	Особенности работы в NoSQL СУБД. Взаимодействие с приграммным приложением	полнота и содержательность ответа умение приводить примеры умение самостоятельно находить решение поставленных задач	ЛЗ – лабораторные задания (3-4), ПЗ – практические задания (2)
В. графовыми и документо-ориентированными СУБД	СУБД MongoDB и Neo4j	полнота и содержательность ответа умение приводить примеры умение самостоятельно находить решение поставленных задач	ЛЗ – лабораторные задания (3-4), ПЗ – практические задания (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).

84-100 баллов (оценка «отлично»),

67-83 баллов (оценка «хорошо»),

50-66 баллов (оценка «удовлетворительно»),

0-49 баллов (оценка «неудовлетворительно»).

**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

***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.

### **Вопросы к экзамену**

- 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) Документо-ориентированные базы данных
- 11) Возможности СУБД MongoDB
- 12) Работа с документо-ориентированными БД на языке JSON

***Экзаменационное задание включает два вопроса – два теоретических вопроса и одно практико-ориентированное задание из числа приведенных ниже лабораторных заданий.***

#### *Критерии оценивания:*

- 84-100 баллов (оценка «отлично») – 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;

- 67-83 баллов (оценка «хорошо») – 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;

- 50-66 баллов (оценка «удовлетворительно») – 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 (оценка «неудовлетворительно») – 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.

### **Задания для опроса**

Вариант 1

Понятие Больших данных.

Особенности сбора, хранения, обработки и анализа Больших данных

Требования к распределенным информационным системам

Вариант 2

Средства построения распределенных информационных систем

Технология Map-Reduce

Система Apache Hadoop

Вариант 3

Базы данных NoSQL. Особенности, классификация

Возможности NoSQL-баз данных по обеспечению целостности, доступности скорости обработки информации. CAP-теорема.

Способы репликации и кластеризации баз данных

Вариант 4

Документо-ориентированные базы данных

Возможности СУБД MongoDB

Работа с документо-ориентированными БД на языке JSON

*Критерии оценивания (для каждого варианта):*

18-20 б. – ответы на все три вопроса варианта даны верно;

16-17 б. – один ответ из 3-х с неточностями;

13-15 б. – 2 ответа из 3-х с неточностями;

9-12 б. – 3 ответа с неточностями;

5-8 б. – нет ответа на один вопрос из 3-х;

1-4 б. – нет ответа на два вопроса из 3-х.

**Максимальное количество баллов за опрос – 20.**

### **Лабораторные задания**

Лабораторное задание №1

Средства построения распределенных информационных систем для BigData. Обзор возможностей

Лабораторное задание №2

Изучение и конфигурирование программного комплекса Apache Hadoop. Размещения набора данных по заданной тематике. Построение поисковых запросов на языке Pig

Лабораторное задание №3

Изучение основных классов NoSQL СУБД, графовых, мультиконочных, документо-ориентированных, типа "имя=значение". Проектирование и разработка графовой базы данных в СУБД Neo4j на заданную тему. Поисковые запросы на языке Cypher

Лабораторное задание №4

Работа с MongoDB. Запросы на выборку и модификацию. Использование драйверов. Настройка фрагментации. Использование Map-Reduce

*Критерии оценивания (для каждого задания):*

8-10 б. – задание выполнено верно;

5-7 б. – при выполнении задания были допущены неточности, не влияющие на результат;

3-4 б. – при выполнении задания были допущены ошибки;

1-2 б. – при выполнении задания были допущены существенные ошибки.

**Максимальное количество баллов за все лабораторные задания – 40 (4 задания по 10 баллов).**



## Практические задания

### Практическое задание №1

Изучение основных классов NoSQL СУБД, графовых, мультиколоночных, документо-ориентированных, типа "имя=значение". Проектирование и разработка графовой базы данных в СУБД Neo4j на заданную тему. Поисковые запросы на языке Cypher

### Практическое задание №2

Работа с MongoDB. Запросы на выборку и модификацию. Использование драйверов. Настройка фрагментации. Использование Map-Reduce

*Критерии оценивания (для каждого задания):*

15-20 б. – задание выполнено верно;

9-14б.– при выполнении задания были допущены неточности, не влияющие на результат;

5-8 б. – при выполнении задания были допущены ошибки;

1-4 б. – при выполнении задания были допущены существенные ошибки.

**Максимальное количество баллов за все практические задания – 40 (2 задания по 20 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, экзамена.

Зачет проводится по окончании теоретического обучения до начала экзаменационной сессии. 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.

Экзамен проводится по расписанию промежуточной аттестации в письменном виде. Количество вопросов в экзаменационном задании – 3 (два теоретических вопроса и одно практико-ориентированное задание). Проверка ответов и объявление результатов производится в день экзамена. Результаты аттестации заносятся в экзаменационную 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;
- практические занятия.

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

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

При подготовке к лабораторным и практическим занятиям каждый студент должен:

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

В процессе подготовки к лабораторным и практическим занятиям студенты могут воспользоваться консультациями преподавателя.

Вопросы, не рассмотренные на лекциях, лабораторных и практических занятиях, должны быть изучены студентами в ходе самостоятельной работы. Контроль самостоятельной работы студентов над учебной программой курса осуществляется в ходе занятий методом опроса и выполнения лабораторных и практических 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 и практическому 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.