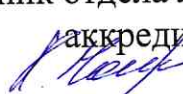


Документ подписан простой электронной подписью  
Информация о владельце:  
ФИО: Макаренко Елена Николаевна  
Должность: Профессор  
Дата подписания: 10.04.2021 17:32:24  
Уникальный программный ключ:  
c098bc0c1041cb2a4cf926cf171d6715d99a6ae00adc8e27b55cbe1e2dbd7c78

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

УТВЕРЖДАЮ

Начальник отдела лицензирования и  
аккредитации



Чаленко К.Н.

« 04 » апрель 2020 г.

**Рабочая программа дисциплины**

**Современные методы распределенного хранения и обработки данных**  
по профессионально-образовательной программе направление 01.03.05 "Статистика"  
профиль 01.03.05.01 "Анализ больших данных"

Для набора 2020 года


Квалификация  
Бакалавр


**КАФЕДРА Информационные технологии и защита информации****Распределение часов дисциплины по семестрам**

Семестр (<Курс>.<Семестр на курсе>)	6 (3.2)		Итого	
	уп	рп	уп	рп
Неделя	16			
Вид занятий	уп	рп	уп	рп
Лабораторные	32	32	32	32
Итого ауд.	32	32	32	32
Контактная работа	32	32	32	32
Сам. работа	76	76	76	76
Итого	108	108	108	108

**ОСНОВАНИЕ**

Учебный план утвержден учёным советом вуза от 25.02.2020 протокол № 8.

Программу составил(и): к.э.н., доцент Рутга Н.А. 

Зав. кафедрой: к.э.н., доцент Е.В. Ефимова 

Методическим советом направления: к.э.н., доцент Кислая И.А. 

<b>1. ЦЕЛИ ОСВОЕНИЯ ДИСЦИПЛИНЫ</b>	
1.1	Изучение современных методов распределенного хранения и обработки данных и формирование практических навыков применения информационных технологий для решения профессиональных задач

<b>2. ТРЕБОВАНИЯ К РЕЗУЛЬТАТАМ ОСВОЕНИЯ ДИСЦИПЛИНЫ</b>	
ПК-1:	способностью подбирать, анализировать и обобщать информацию, содержащуюся в методических документах и научно-технической литературе, в том числе на иностранном языке
ПК-11:	способностью обеспечивать сохранность и конфиденциальность индивидуальных данных и другой статистической информации

<b>В результате освоения дисциплины обучающийся должен:</b>	
<b>Знать:</b>	понятия и методы распределенного хранения и обработки данных с применением информации, содержащейся в методических документах и научно-технической литературе, в том числе на иностранном языке методы сохранения и конфиденциальности индивидуальных данных и другой статистической информации
<b>Уметь:</b>	выбирать методы распределенного хранения и обработки данных для решения профессиональных задач выбирать методы сохранения и конфиденциальности индивидуальных данных и другой статистической информации
<b>Владеть:</b>	навыками выбора и применения различных методов распределенного хранения и обработки данных для решения профессиональных задач
<b>Методы:</b>	сохранности и конфиденциальности индивидуальных данных и другой статистической информации

<b>3. СТРУКТУРА И СОДЕРЖАНИЕ ДИСЦИПЛИНЫ</b>					
Код задания	Наименование разделов и тем /внд занятия/	Семестр / Курс	Часов	Компетен-ции	Литература
1.1	<b>Раздел 1. Распределенная обработка и хранение данных: SQL Server, Dedeator</b> Лабораторная работа 1. Основы работы в Microsoft SQL Server: Создание базы данных: таблиц, установка связей с помощью Давянье Даврагаш, ввод данных, отключение и подключение базы данных //Лаб/	6	4	ПК-11 ПК-1	Л1.4 Л1.7Л2.6
1.2	Лабораторная работа 2. Создание запросов и фильтров в Microsoft SQL Server. Обработка данных с помощью оператора SELECT-SQL с помощью утилиты ISQL. Запрос с группировкой и применение агрегатных функций WHERE. Запрос с вычисляемыми полями	6	4	ПК-11 ПК-1	Л1.5Л2.6
1.3	Лабораторная работа 3. Создание и использование хранимых процедур в Microsoft SQL Server. Оператор CREATE PROCEDURE. Создание процедур без параметров, с входными параметром, с выходными параметрами и выходными параметром и т.д. Хранимые функции Реализация системы защиты в MS SQL Server //Лаб/	6	4	ПК-11 ПК-1	Л1.5 Л1.7Л2.6
1.4	Лабораторная работа 4. Разработка хранилища данных в Dedeator: Создание хранилища Dedeator Warehouse Редактор метаданных: атрибут пропееса, атрибут измерения Наполнение хранилища Полноценное хранилища данных Создание многомерных отчетов (OLAP-кубов) и кросс-диаграмм и их анализа //Лаб/	6	4	ПК-11 ПК-1	Л1.2Л2.1

1.5	Вопросы для самостоятельного изучения: Основное понятие распределенной обработки данных Условия и требования к распределенной обработке данных Архитектура распределенной обработки данных: Архитектура «файл-сервер» Архитектура «выделенный сервер базы данных» Архитектура «сервер приложений» Хранилища данных: свюкства, данные, источник данных Витрина данных Компоненты хранилища Выбор метода реализации хранилищ данных Методы защиты данных в распределенных базах	6	30	ПК-11 ПК-1	Л1.4 Л1.5 Л2.1 Л1.7Л2.2 Л2.6
	<b>Раздел 2. Обязательные семинары в обработке данных. Обработка данных в ArcSde</b>	6	6	ПК-11 ПК-1	Л1.1 Л1.6Л2.5

2.1	Лабораторная работа 5. Обязательные технологии в обработке данных Работа с онлайн-сервисом Google Таблицы. Создание опроса с помощью Google Форм. Создание теста и обработка результатов Облачные хранилище Teambox от Mail.ru //Лаб/	6	6	ПК-11 ПК-1	Л1.3 Л1.4Л2.4
2.2	Лабораторная работа 6. Платформа ArcSde Принципы распределенной реализации MapReduce на кластерных системах. Платформа ArcSde Hadoop. Интерфейсы прикладного программирования и реализации программ для Hadoop. Локальная отладка и запуск программ на кластере. Приемы и стратегии реализации MapReduce-приорграмм. Изучение структуры кодов //Лаб/	6	6	ПК-11 ПК-1	Л1.1 Л1.6Л2.3
2.3	Лабораторная работа 7. Big Data as-a-Service Cloud Big Data. Масштабируемый PaaS-сервис для анализа больших данных на базе ArcSde Hadoop, ArcSde Spark //Лаб/	6	4	ПК-11 ПК-1	Л1.1 Л1.6Л2.3
2.4	Вопросы для самостоятельного изучения: Облачные хранилища Системы управления потоками данных. Системы хранения больших данных. Платформа больших данных. Обработка данных в реальном времени. Система управления большими данными. Аналитические платформы. Состав и возможности программного комплекса ArcSde Hadoop. Задачи поисковых запросов для Hadoop. Принципы работы ArcSde Spark Компоненты экосистемы Hadoop. /Ср/	6	46	ПК-11 ПК-1	Л1.1 Л1.3 Л1.4 Л1.4 Л1.5 Л1.6 Л1.6Л2.3 Л2.4 Л2.5
2.5	/Зачет/	6	0	ПК-11 ПК-1	Л1.1 Л1.2 Л1.3 Л1.4 Л1.5 Л1.6 Л1.7Л2.2 Л2.3 Л2.1 Л2.4 Л2.5 Л2.6

<b>4. ФОНД ОЦЕНОЧНЫХ СРЕДСТВ</b>	
Структура и содержание фонда оценочных средств для проведения текущей и промежуточной аттестации представлены в Приложении 1 к рабочей программе дисциплины.	

<b>5. УЧЕБНО-МЕТОДИЧЕСКОЕ И ИНФОРМАЦИОННОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ</b>			
<b>5.1. Основная литература</b>			
Авторы, составители	Заглавие	Издательство, год	Копий-во



	Авторы, составители	Заглавие	Издательство, год	Копия-во
Л1.1	Губарева В. В., Савурчич С. А., Чистяков Н. А.	Введение в облачные вычисления и технологии: учебное пособие	Новосибирск: Новосибирский госуниверситетский технический университет, 2013	<a href="http://bibliobd.uib.niindex.ru/">http://bibliobd.uib.niindex.ru/</a> page=book&id=228962 неограниченный доступ для зарегистрированных пользователей
Л1.2		Хранилища данных	Москва: Национальный Открытый Университет «ИНТУИТ», 2014	<a href="http://bibliobd.uib.niindex.ru/">http://bibliobd.uib.niindex.ru/</a> page=book&id=237117 неограниченный доступ для зарегистрированных пользователей
Л1.3	Волкова, Т. В., Насежкина, Л. Ф.	Разработка систем распределенной обработки данных: учебно-методическое пособие	Оренбург: Оренбургский государственный университет, ЭБС АСВ, 2012	<a href="http://www.irbookshop.ru/">http://www.irbookshop.ru/</a> u/30127.html неограниченный доступ для зарегистрированных пользователей
Л1.4	Косяков, М. С.	Введение в распределенные вычисления	Санкт-Петербург: Университет ИТМО, 2014	<a href="http://www.irbookshop.ru/">http://www.irbookshop.ru/</a> u/65816.html неограниченный доступ для зарегистрированных пользователей
Л1.5	Зудилова, Т. В., Шмелева, Г. Ю.	Создание запросов в Microsoft SQL Server 2008	Санкт-Петербург: Университет ИТМО, 2013	<a href="http://www.irbookshop.ru/">http://www.irbookshop.ru/</a> u/68136.html неограниченный доступ для зарегистрированных пользователей
Л1.6	Званитирова, Л. Ф.	Облачные вычисления: учебное пособие	Саратов: Ай Пи Ар Медиа, 2019	<a href="http://www.irbookshop.ru/">http://www.irbookshop.ru/</a> u/85804.html неограниченный доступ для зарегистрированных пользователей
Л1.7		Разработка и защита баз данных в Microsoft SQL Server 2005: учебное пособие для СПО	Саратов: Профобразование, 2019	<a href="http://www.irbookshop.ru/">http://www.irbookshop.ru/</a> u/86207.html неограниченный доступ для зарегистрированных пользователей
<b>5.2. Дополнительная литература</b>				
	Авторы, составители	Заглавие	Издательство, год	Копия-во
Л2.1	Парфенов, Ю. П.	Постреляционные хранилища данных: учебное пособие	Екатеринбург: Уральский федеральный университет, ЭБС АСВ, 2016	<a href="http://www.irbookshop.ru/">http://www.irbookshop.ru/</a> u/68372.html неограниченный доступ для зарегистрированных пользователей
Л2.2		Информационная безопасность: журнал	Москва: ГРОТЕК, 2014	<a href="http://bibliobd.uib.niindex.ru/">http://bibliobd.uib.niindex.ru/</a> page=book&id=230502 неограниченный доступ для зарегистрированных пользователей
Л2.3	Воронова, Л. И., Воронов, В. И.	Big Data. Методы и средства анализа: учебное пособие	Москва: Московский технический университет связи и информатики, 2016	<a href="http://www.irbookshop.ru/">http://www.irbookshop.ru/</a> u/61463.html неограниченный доступ для зарегистрированных пользователей

	Авторы, составители	Заглавие	Издательство, год	Копия-во
Л2.4	Воронова, В. И., Воронова, Л. И., Усачев, В. А.	Data Mining - технологии обработки больших данных: учебное пособие	Москва, Московский технический университет связи и информатики, 2018	<a href="http://www.irbookshop.ru/">http://www.irbookshop.ru/</a> u/81324.html неограниченный доступ для зарегистрированных пользователей
Л2.5	Балуев, Д.	Серветы приложений Google	Москва: Альпина Паблишер, 2019	<a href="http://www.irbookshop.ru/">http://www.irbookshop.ru/</a> u/82484.html неограниченный доступ для зарегистрированных пользователей
Л2.6	Митин А. И.	Работа с базами данных Microsoft SQL Server: сценарии практических занятий: практикум	Москва: Берлин: Директ-Медиа, 2020	<a href="http://bibliobd.uib.niindex.ru/">http://bibliobd.uib.niindex.ru/</a> page=book&id=571169 неограниченный доступ для зарегистрированных пользователей

### 5.3 Профессиональные базы данных и информационные справочные системы

Консультант плюс				
Электронно-библиотечная система «Университетская библиотека онлайн»	<a href="http://bibliobd.uib.niindex.ru/">http://bibliobd.uib.niindex.ru/</a>			
Электронно-библиотечная система «IPK Books»	<a href="http://www.irbookshop.ru/">http://www.irbookshop.ru/</a>			
Официальная документация <a href="https://docs.microsoft.com/ru-ru/sql/tds/big-data-cluster/big-data-cluster-overview?view=sql-server-ver15">https://docs.microsoft.com/ru-ru/sql/tds/big-data-cluster/big-data-cluster-overview?view=sql-server-ver15</a>				
Официальная документация <a href="https://hadoop.apache.org/docs/current/">https://hadoop.apache.org/docs/current/</a>				

### 5.4 Перечень программного обеспечения

MS SQL SERVER				
Debian				
Apache Hadoop				
Google таблицы ( <a href="https://www.google.ru/mi/gu/sheets/about/">https://www.google.ru/mi/gu/sheets/about/</a> )				

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

При необходимости по заявлению обучающихся 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, комплектованы необходимой специализированной учебной мебелью и техническими средствами обучения. Лабораторные занятия проводятся в компьютерных классах, рабочие места в которых оборудованы необходимыми лицензионными программами 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.

## Приложение 1

### 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:

ЗУИ, составляющие компетенцию	Показатели оценивания	Критерии оценивания	Средства оценивания
ПК-1: способностью подбирать, анализировать и обобщать информацию, содержащуюся в методических документах и научно-технической литературе, в том числе на иностранном языке			
3. понятия и методы распределенного хранения и обработки данных с применением информации, содержащейся в методических документах и научно-технической литературе, в том числе на иностранном языке	изучает методические документы, научно-техническую литературу для освоения методов распределенного хранения и обработки данных, для подготовки к зачету и устному опросу	полнота и содержательность ответа на зачете, устном опросе, соответствие содержания материала, содержащегося в изученной литературе	УО (1-15, 23-29) 3 (1-30)
У. выбирать методы распределенного хранения и обработки данных для решения профессиональных задач	решает практико-ориентированные лабораторные задания: разрабатывает хранилища данных с использованием программного средства Deducator	правильность построения хранилища данных с использованием программного средства Deducator	Л3 (1-4) ПО3 (1,2)
В. владения навыками выбора и применения различных методов распределенного хранения и обработки данных для решения профессиональных задач	создает Google Таблицы, формы различные подходы к анализу кодов MarReduce при выполнении лабораторных и практико-ориентированных	обоснованность применения различных функций в Google Таблицах, форм правильность и обоснованность различных подходов к анализу кодов	Л3 (5-7) ПО3 (3,4)

задачи	заданий	МарReduce	
ПК-11: способностью обеспечивать сохранность и конфиденциальность индивидуальных данных и другой статистической информации			
3. методы сохранности и конфиденциальности индивидуальных данных и другой статистической информации	изучает литературу, содержащую информацию о разных методах сохранности и конфиденциальности индивидуальных данных и другой статистической информации, для подготовки к зачету и устному опросу	полнота и содержательность ответа на зачете, устном опросе, соответствие содержания материала, содержащегося в изученной литературе	УО (16-22, 30-32) 3 (13-30)
У. - выбирать методы сохранности и конфиденциальности индивидуальных данных и другой статистической информации	решает практико-ориентированные лабораторные задания: создает таблицы, строит запросы и ответы Microsoft SQL Server	правильность создания таблиц, построения запросов и ответов Microsoft SQL Server	Л3-3 Л3-4 Л3-5 ПО3 (1,2,3)
В. Навыками применения методов сохранности и конфиденциальности индивидуальных данных и другой статистической информации	решает практико-ориентированные лабораторные задания различными способами, которые применяются при построении запросов и ответов Microsoft SQL Server	правильность и обоснованность выбора методов построения запросов и ответов Microsoft SQL Server	Л3-3 Л3-4 Л3-5 ПО3 (1,2,3)

3 – вопросы к зачету, Л3 – лабораторные задания, ПО3-практико-ориентированное задание к зачету; УО-устный опрос

##### 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



В разделе приводятся типовые варианты оценочных средств: вопросы к зачету вопросы для устного опроса, практические задания.

**Вопросы к зачету по Дисциплине**

1. Основные понятия распределённой обработки данных
2. Условия и требования к распределённой обработке данных
3. Основные принципы создания и функционирования распределённых баз данных
4. Архитектура распределённой обработки данных: Архитектура «файл-сервер»
5. Архитектура «выделенный сервер базы данных»
6. Архитектура «сервер приложений»
7. Технологии объектного связывания данных
8. Хранилища данных: свойства, данные, источник данных
9. Витрины данных
10. Компоненты хранилища
11. Выбор метода реализации хранилищ данных
12. Оптимизация хранилищ данных
13. Организация защиты конфиденциальной информации
14. Методы защиты данных в распределённых базах
15. Уязвимости и угрозы безопасности данных
16. Средства защиты баз данных
17. Облачные системы обработки данных.
18. Облачные хранилища
19. Системы управления потоками данных.
20. Системы хранения больших данных.
21. Платформы больших данных.
22. Обработка данных в реальном времени.
23. Системы управления большими данными.
24. Аналитические платформы.
25. Состав и возможности программного комплекса Apache Hadoop.
26. Языки поисковых запросов для Hadoop.
27. Принципы работы Apache Spark.
28. Компоненты экосистемы Hadoop.
29. MapReduce — модель распределённых вычислений
30. Raas-сервис для анализа больших данных на базе Apache Hadoop

**Практико-ориентированные задания к зачету**

**Задание 1 (Microsoft SQL Server)**

- 1) Создать базу данных

**БД «Доступление лекарственных средств»**

Код лекарства	Код лекарства	Код поставщика
Название лекарства	Код поставщика	Сокращённое название
Показания к применению	Дата поставки	Полное название
Единица измерения	Цена за единицу	Корrespondентский адрес
Назначение провозводителя	Колличество	Телефон
	Код поступления	ФИО руководителя

- 2) Разработать следующие запросы:
  1. Вывести список лекарств с указанным показанием к применению.
  2. Вывести список дат поставок, по которым продано больше указанного числа одноименного лекарства.
  3. Вывести дату поставки, сумму, ФИО руководителя от поставщика и название лекарства по коду поступления больше указанного числа.
  4. Вывести список лекарств и единицы измерения для лекарств с количеством в упаковке больше указанного числа или кодом лекарства меньше определённого значения

**Задание 2 (Deductor)**

Разработать хранилище данных

«Учет выполненных ремонтных работ»

Код прибора в ремонте	Код прибора	Код мастера
Название прибора	Код мастера	Фамилия мастера
Тип прибора	ФИО владельца прибора	Имя мастера
Дата производства	Дата приема в ремонт	Отчество мастера
	Вид поломки	Разряд мастера
	Стоимость ремонта	Дата приема на работу
	Код ремонта	

**Задание 3 (Google Таблицы)**

1. Разработайте таблицу-прайс. Ввести исходные данные (5-6 строк)
2. Применить следующие функции:
  - IF.
  - IFERROR
  - JOIN
3. Сгенерировать QR-код
4. Привязать таблицу к Google-форме (предварительно разработать)
5. Провести защиту диапазона ячеек

#### Задание 4 (MapReduce)

Предположим, у нас есть данные о сотрудниках в четырех разных файлах — A, B, C и D. Предположим также, что во всех четырех файлах есть дубликаты записей о сотрудниках из-за повторного импорта данных о сотрудниках из всех таблиц базы данных. (См рисунок).

name, salary	name, salary	name, salary	name, salary
saiish, 26000	gorral, 50000	saiish, 26000	saiish, 26000
Krishna, 25000	Krishna, 25000	Kiran, 45000	Krishna, 25000
Saishk, 15000	Saishk, 15000	Saishk, 15000	Saishk, 15000
Rajd, 10000	Rajd, 10000	Rajd, 10000	Rajd, 10000

**Фаза Map** обрабатывает каждый входной файл и предоставляет данные о сотруднике в парах ключ-значение (<k, v>; <emp name, salary>).

1. Изобразите иллюстрацией этот процесс

```
<k: employee name, v: salary>  
Max= the salary of an first employee. Treated as max salary
```

```
if(v(second employee).salary > Max){  
    Max = v(salary);  
}  
else{  
    Continue checking;  
}
```

#### Критерии оценивания:

- 50-100 баллов (зачет) – изложенный материал фактически верен, наличие глубоких исчерпывающих знаний; правильные, уверенные действия по применению полученных знаний на практике; грамотное и логически стройное изложение материала при ответе; практико-ориентированное задание выполнено правильно и прокомментировано; наличие твердых и достаточно полных знаний, правильные действия по применению знаний на практике, четкое изложение материала, допускаются отдельные логические и стилистические погрешности, неуверенность и неточность ответов на дополнительные вопросы и наводящие вопросы; практико-ориентированное задание выполнено правильно, но не прокомментировано; при неполном ответе на вопросы; затрудняется ответить на дополнительные вопросы; практико-

ориентированное задание выполнено с ошибками и отсутствующим комментарием.

30 баллов за 1 теоретический вопрос

30 баллов за 2 теоретический вопрос

40 баллов за выполненное практико-ориентированное задание

- 0-49 баллов (незачет) – ответы не связаны с вопросами, наличие грубых ошибок в ответе, непонимание сути вопроса, отсутствие ответов на дополнительные и наводящие вопросы; практико-ориентированное задание не выполнено.

#### Лабораторные задания

##### Раздел 1. Распределенная обработка и хранение данных: SQL Server, Deductor

Лабораторное задание 1 Основы работы в Microsoft SQL Server (5 баллов).

Создание базы данных: таблицы, установка связей с помощью Database Diagram, ввод данных, отключение и подключение базы данных

Лабораторное задание 2. Создание запросов и фильтров в Microsoft SQL Server. (10 баллов)

Обработка данных с помощью оператора SELECT-SQL с помощью утилиты ISQL.

Задание условий отбора записей WHERE.

Запрос с вычисляемым полем  
Запрос с группировкой и применение агрегатных функций

Лабораторное задание 3. Создание и использование хранимых процедур в Microsoft SQL Server. (9 баллов)

Оператор CREATE PROCEDURE.  
Создание процедур без параметров, с входным параметром, с входными параметрами и выходным параметром и т.д.

Хранимые функции  
Реализация системы защиты в MS SQL Server

Лабораторное задание 4. Разработка хранилища данных в Deductor. (10 баллов)

Создание хранилища Deductor Warehouse.  
Редактор метаданных: атрибут процесса, атрибут измерения



- Наполнение хранилища.
- Подключение хранилища данных.
- Создание многомерных отчетов (OLAP-кубов) и кросс-диаграмм и их анализ
- Максимальное количество баллов по лабораторным заданиям – 34

## Раздел 2. Облачные сервисы в обработке данных. Обработка данных в

### Арасче

- Лабораторное задание 5. Облачные технологии в обработке данных (10 баллов).
- Работа с табличным онлайн процессором Google Таблицы. Создание опроса с помощью Google Форм. Создание теста и обработка результатов
- Облачное хранилище Teambox от Mail.ru

- Лабораторное задание 6. Платформа Арасче (20 баллов)
- Принципы распределенной реализации MapReduce на кластерных системах.
- Платформа Арасче Hadoop.
- Интерфейсы прикладного программирования и реализации программ для Hadoop.
- Локальная отладка и запуск программ на кластере. Приемы и стратегии реализации MapReduce-программ.
- Изучение структуры кодов

Лабораторное задание 7. Big Data as-a-Service данных (10 баллов).

- Cloud Big Data. Масштабируемый PaaS-сервис для анализа больших данных на базе Арасче Hadoop, Арасче Spark

Максимальное количество баллов по лабораторным заданиям – 40

### Перечень вопросов для устного опроса

## Раздел 1. Распределенная обработка и хранение данных: SQL Server, Reductor

1. Назовите основные проводники, браузеры, редакторы, визуальные конструкторы и мастера среды SQL Server Management Studio.
2. Что такое источники и поставщики данных?
3. Что такое индекс?
4. Дайте определение типа данных.
5. В чем заключается понятие сущности, атрибута, кортежа?
6. Какие виды ключей таблиц вы знаете?
7. Для чего используются таблицы в БД?

8. Какие типы данных полей есть в SQL Server?
9. Как можно создать и заполнить таблицы?
10. Что такое целостность данных?
11. Что такое диаграмма базы данных?
12. Как создаются запросы?
13. Для чего используются фильтры?
14. В чем состоит процесс выполнения вычислений при помощи оператора SELECT?

15. Какие встроены функции могут использоваться в операторе SELECT?
16. Что такое хранилище данных?
17. Что такое виртуальное и физическое хранилище данных?
18. Что такое витрина данных?
19. Из чего состоит хранилище данных?
20. Какие потоки данных имеют в хранилище данных?
21. Какие есть приемы оптимизации хранения данных?
22. Охарактеризуйте угрозы безопасности информации, связанные с централизацией в базах данных большого количества разнообразной информации

### Критерии оценивания:

- Для каждого вопроса:
  - 1 балл дан полный, развернутый ответ на поставленный вопрос, изложение материала при ответе – грамотное и логически стройное;
  - 0 баллов – обучающийся не владеет материалом по заданному вопросу.

Максимальное количество баллов – 16

## Раздел 2. Облачные сервисы в обработке данных. Обработка данных в

### Арасче

23. В чем заключается сущность онлайн сервиса электронных таблиц Google, для чего предназначена Арасче Hadoop?
24. Для чего предназначена Арасче Hadoop?
25. Из каких компонентов состоит Hadoop?
26. В чем заключается парадигма MapReduce?
27. В чем заключается фаза map?
28. В чем заключается фаза reduce?
29. Какие преимущества дает использование парадигмы MapReduce при обработке «больших данных» (big data)?
30. Что такое облачное хранилище?
31. Как работает облачное хранилище?
32. Преимущества облачного хранилища



**Критерии оценивания:**

Для каждого вопроса:

- 1 балл дан полный, развернутый ответ на поставленный вопрос, изложение материала при ответе – грамотное и логически стройное;
- 0 баллов – обучающийся не владеет материалом по заданному вопросу.

Максимальное количество баллов – 10

**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 окончании теоретического обучения до начала экзаменационной сессии. Количество вопросов в задании – 3 (2 теоретических вопроса и 1 практико-ориентированное задание). Проверка ответов и 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:

– лабораторные работы.

В ходе лабораторных работ углубляются навыки практического применения современных методов обработки и хранения данных с помощью информационных технологий, а также навыки анализа и применения информации, содержащейся в методических документах и научно-технической литературе.

При подготовке к лабораторным занятиям каждый студент должен:

– изучить рекомендованную учебную литературу;

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

Теоретические вопросы должны быть изучены студентами в ходе самостоятельной работы. Контроль самостоятельной работы студентов над учебной программой курса осуществляется методом устного опроса или выполнения практико-ориентированных заданий. В ходе самостоятельной работы каждый студент обязан прочитать основную и по возможности дополнительную литературу по изучаемой теме. Выделить непонятные термины, найти их значение в литературе.

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.