



МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ  
РОССИЙСКОЙ ФЕДЕРАЦИИ

Федеральное государственное бюджетное образовательное учреждение  
высшего образования  
«Магнитогорский государственный технический университет им. Г.И. Носова»



СВЕРЖДАЮ  
Директор ИЭиАС  
И. Лукьянов

26.02.2020 г.

**РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ (МОДУЛЯ)**

***ОСНОВЫ DATA ИНЖИНИРИНГА***

Направление подготовки (специальность)

10.05.03 ИНФОРМАЦИОННАЯ БЕЗОПАСНОСТЬ АВТОМАТИЗИРОВАННЫХ СИСТЕМ

Направленность (профиль/специализация) программы

10.05.03 специализация N 7 "Обеспечение информационной безопасности распределенных информационных систем";

Уровень высшего образования - специалитет

Форма обучения  
очная

Институт/ факультет	Институт энергетики и автоматизированных систем
Кафедра	Информатики и информационной безопасности
Курс	2
Семестр	3

Магнитогорск  
2020 год

Рабочая программа составлена на основе ФГОС ВО по специальности 10.05.03  
ИНФОРМАЦИОННАЯ БЕЗОПАСНОСТЬ АВТОМАТИЗИРОВАННЫХ СИСТЕМ  
(приказ Минобрнауки России от 01.12.2016 г. № 1509)

Рабочая программа рассмотрена и одобрена на заседании кафедры Информатики и  
информационной безопасности  
18.06.2020, протокол № 6

Зав. кафедрой  И.И. Баранкова

Рабочая программа одобрена методической комиссией ИЭиАС  
26.02.2020 г. протокол № 5

Председатель  С.И. Лукьянов

Рабочая программа составлена:

доцент кафедры ИиИБ, канд. техн. наук  Л.Л. Демиденко

Рецензент:

начальник УИТ и АСУ, канд. техн. наук  К.А. Рубан

## Лист актуализации рабочей программы

---

Рабочая программа пересмотрена, обсуждена и одобрена для реализации в 2021 - 2022 учебном году на заседании кафедры Информатики и информационной безопасности

Протокол от \_\_\_\_\_ 20\_\_ г. № \_\_\_\_  
Зав. кафедрой \_\_\_\_\_ И.И. Баранкова

---

Рабочая программа пересмотрена, обсуждена и одобрена для реализации в 2022 - 2023 учебном году на заседании кафедры Информатики и информационной безопасности

Протокол от \_\_\_\_\_ 20\_\_ г. № \_\_\_\_  
Зав. кафедрой \_\_\_\_\_ И.И. Баранкова

---

Рабочая программа пересмотрена, обсуждена и одобрена для реализации в 2023 - 2024 учебном году на заседании кафедры Информатики и информационной безопасности

Протокол от \_\_\_\_\_ 20\_\_ г. № \_\_\_\_  
Зав. кафедрой \_\_\_\_\_ И.И. Баранкова

---

Рабочая программа пересмотрена, обсуждена и одобрена для реализации в 2024 - 2025 учебном году на заседании кафедры Информатики и информационной безопасности

Протокол от \_\_\_\_\_ 20\_\_ г. № \_\_\_\_  
Зав. кафедрой \_\_\_\_\_ И.И. Баранкова

---

Рабочая программа пересмотрена, обсуждена и одобрена для реализации в 2025 - 2026 учебном году на заседании кафедры Информатики и информационной безопасности

Протокол от \_\_\_\_\_ 20\_\_ г. № \_\_\_\_  
Зав. кафедрой \_\_\_\_\_ И.И. Баранкова

### 1 Цели освоения дисциплины (модуля)

Целью изучения дисциплины «Основы Data Engineering» является формирование у обучающихся компетенций в области современных технологий обработки и анализа Big Data в соответствии с требованиями ФГОС ВО для специальности 10.05.03 «Информационная безопасность автоматизированных систем».

### 2 Место дисциплины (модуля) в структуре образовательной программы

Дисциплина Основы Data инжиниринга входит в базовую часть учебного плана образовательной программы.

Для изучения дисциплины необходимы знания (умения, владения), сформированные в результате изучения дисциплин/ практик:

Введение в специальность

Информатика

Теория информации

Основы безопасности цифрового общества

Знания (умения, владения), полученные при изучении данной дисциплины будут необходимы для изучения дисциплин/практик:

Безопасность систем баз данных

Информационные технологии. Базы данных

Техническая защита информации

Защита электронного документооборота

### 3 Компетенции обучающегося, формируемые в результате освоения дисциплины (модуля) и планируемые результаты обучения

В результате освоения дисциплины (модуля) «Основы Data инжиниринга» обучающийся должен обладать следующими компетенциями:

Структурный элемент компетенции	Планируемые результаты обучения
ОПК-3 способностью применять языки, системы и инструментальные средства программирования в профессиональной деятельности	
Знать	<ul style="list-style-type: none"><li>- базовые понятия и терминологию Big Data;</li><li>- возможности инструментальных средств программирования для обработки Big Data;</li><li>- основные алгоритмы и структуры для работы с Big Data;</li><li>- основные понятия распределённых систем.</li></ul>
Уметь	<ul style="list-style-type: none"><li>- применять инструментальные средства программирования для обработки Big Date;</li><li>- выявлять скрытые аномалии в данных</li></ul>
Владеть	<ul style="list-style-type: none"><li>- современными технологиями обслуживания Big Data;</li><li>- алгоритмами и методами для обработки Big Data;</li><li>- инструментальными средствами программирования для обработки Big Data.</li></ul>

ОПК-4 способностью понимать значение информации в развитии современного общества, применять достижения современных информационных технологий для поиска информации в компьютерных системах, сетях, библиотечных фондах	
Знать	<ul style="list-style-type: none"> <li>- основные направления развития, значение Big Data и их роль в современном обществе;</li> <li>- основные возможности ИКТ для поиска Big Data из различных источников информации.</li> </ul>
Уметь	<ul style="list-style-type: none"> <li>- организовывать поиск Big Data в компьютерных системах, сетях и библиотечных фондах;</li> <li>- автоматизировать работу с большими массивами данных, получать данные из внешних источников.</li> </ul>
Владеть	<ul style="list-style-type: none"> <li>- методами поиска, хранения и обработки Big Data из различных источников баз данных, библиотечных фондов;</li> <li>- методами представления Big Data с использованием информационных, компьютерных и сетевых технологий.</li> </ul>
ПК-1 способностью осуществлять поиск, изучение, обобщение и систематизацию научно-технической информации, нормативных и методических материалов в сфере профессиональной деятельности, в том числе на иностранном языке	
Знать	<ul style="list-style-type: none"> <li>- особенности обработки информации в сфере Big Data</li> </ul>
Уметь	<ul style="list-style-type: none"> <li>- обобщать, систематизировать, обеспечивать надежную инфраструктуру полученной научно-технической информации в сфере data инжиниринг.</li> </ul>
Владеть	<ul style="list-style-type: none"> <li>- методами сбора и обобщения научно-технической информации по теоретическим сведениям, проблемам и задачам, решаемым в курсе "Data Инжиниринг", в т.ч. с использованием литературы на иностранном языке.</li> </ul>

#### 4. Структура, объём и содержание дисциплины (модуля)

Общая трудоемкость дисциплины составляет 3 зачетных единиц 108 академических часов, в том числе:

- контактная работа – 73,9 академических часов;
- аудиторная – 72 академических часов;
- внеаудиторная – 1,9 академических часов;
- самостоятельная работа – 34,1 академических часов;

Форма аттестации - зачет

Раздел/ тема дисциплины	Семестр	Аудиторная контактная работа (в академических часах)			Самостоятельная работа	Вид самостоятельной работы	Форма текущего контроля успеваемости и промежуточной аттестации	Код компетенции
		лек.	лаб. зан.	практ. зан.				
1. Основные понятия курса "основы Data Engineering". Технологии обработки Big Data. Отличие от традиционной технологии.								
1.1 Понятие Big Data. Технологии хранения и обработки Big Data. Три определяющих характеристики - основные принципы работы - 4V (volume, velocity, variety, value) с Big Data.	3	2			1	Самостоятельное изучение учебной и научной литературы, работа с материалами образовательного портала и ЭБС. Подготовка реферата.	реферат	ОПК-3, ОПК-4, ПК-1
Итого по разделу		2			1			
2. Системы управления Big data.								
2.1 Современные технологии обслуживания Big Data. Техники и методики анализа, применимые к Big data.	3	1			1	Самостоятельное изучение учебной и научной литературы, работа с материалами образовательного портала и ЭБС. Подготовка к тестированию	Тестирование	ОПК-3, ОПК-4

2.2 Программное обеспечение, применяемое для структурирования и анализа Big Data. Консолидация и хранение данных в едином пространстве и единой нормализованной модели		1			2	Самостоятельное изучение учебной и научной литературы, работа с материалами образовательного портала и ЭБС. Подготовка к тестированию	Тестирование	ОПК-3
Итого по разделу		2			3			
3. Роль языков программирования Python и R в аналитике больших данных.								
3.1 Основные элементы языка Питон. Операции, функции, Преобразование типов данных, строки. Управляющие конструкции	3	4	6		1	Самостоятельное изучение учебной и научной литературы, работа с материалами образовательного портала и ЭБС. Подготовка к контр работе	Контр работа	ОПК-3
3.2 Списки, кортежи, множества. Исключения, словари. ООП – классы, объекты, свойства, методы		4	2		2	Самостоятельное изучение учебной и научной литературы, работа с материалами образовательного портала и ЭБС.	Защита лабораторной работы	ОПК-3
3.3 Чтение файлов. Файлы с разделителями. Извлечение данных из Web –ресурсов. Использование программных интерфейсов. Импорт модулей. Регулярные выражения		2	4		2	Самостоятельное изучение учебной и научной литературы, работа с материалами образовательного портала и ЭБС. Подготовка ИДЗ	ИДЗ	ОПК-3 ОПК-4, ПК-1
3.4 Анализ данных (пакеты scipy, scikit-learn, pandas, numpy). Фильтрация и сортировка, группировка, сложная агрегация		6	2		3	Самостоятельное изучение учебной и научной литературы, работа с материалами образовательного портала и ЭБС. Подготовка к сдаче лаб работы	Защита лабораторной работы	ОПК-3, ОПК-4, ПК-1

3.5	Основы архитектуры хранения и обработки больших данных, виды обработки и масштабирования. Синтаксический анализ записей файла больших размеров с использованием стандартных функций и библиотек. Управление данными, уменьшение размерности.		4	4		4	Самостоятельное изучение учебной и научной литературы, работа с материалами образовательного портала и ЭБС. Подготовка к защите лабораторной работы	Защита лабораторной работы	ОПК-3, ПК-1
3.6	Feature Engineering - пропущенные значения и дублирование. Запуск скриптов. Парсинг динамических элементов средствами Python. Парсинг фото и файлов		2			5,5	Самостоятельное изучение учебной и научной литературы, работа с материалами образовательного портала и ЭБС. Подготовка к защите лабораторной работы.	Защита лабораторной работы	ОПК-4, ПК-1
Итого по разделу			22	1		18,			
4. Методы сбора и обработки данных из сети Internet									
4.1	Основы клиент-серверного взаимодействия. Парсинг API. Парсинг HTML. BeautifulSoup, MongoDB	3	2			0,5	Самостоятельное изучение учебной и научной литературы, работа с материалами образовательного портала и ЭБС. Подготовка реферата	Реферат	ОПК-4, ПК-1
Итого по разделу			2			0,5			
5. Организация big data. Реляционные СУБД. СУБД NoSQL. Кластеризация									
5.1	Подходы к организации big data: реляционных SQL и noSQL. Технология MapReduce. Системы управления базами данных MongoDB в Python	3	2			2	Самостоятельное изучение учебной и научной литературы, работа с материалами образовательного портала и ЭБС. Подготовка к тестированию	Тестирование	ОПК-4, ПК-1
Итого по разделу			2			2			
6. Инструменты для аналитики Big Data. Системы анализа Big Data									

<p>6.1 Культура сбора и источники данных Предобработка и визуализация данных в pandas Улучшение качества работы с данными</p>		0,5			2	<p>Самостоятельное изучение учебной и научной литературы, работа с материалами образовательного портала и ЭБС. Подготовка к тестированию</p>	Тестирование	ОПК-3 ОПК-4, ПК-1
<p>6.2 Использование Hadoop MapReduce – платформы программирования и выполнения распределённых MapReduce-вычислений, с использованием большого количества компьютеров (узлов, nodes), образующих кластер.</p>	3	1,5	6		1,1	<p>Самостоятельное изучение учебной и научной литературы, работа с материалами образовательного портала и ЭБС. Подготовка к защите лабораторной работы</p>	Защита лабораторной работы	ОПК-3, ОПК-4, ПК-1
<p>6.3 Основы работы в Hadoop и MapReduce инструментами экосистемы Hadoop: HDFS, Yarn, Hive, Hue, Flume, Cassandra Инструменты класса Data Discovery. Apache Spark. Apache Kafka.</p>		1	6		2	<p>Самостоятельное изучение учебной и научной литературы, работа с материалами образовательного портала и ЭБС.</p>	Защита лабораторной работы	ОПК-3, ОПК-4, ПК-1
<p>6.4 Работа с облачными платформами: AWS, EMR, Azure Продвинутые подходы в MapReduce</p>		2	2		4	<p>Самостоятельное изучение учебной и научной литературы, работа с материалами образовательного портала и ЭБС.</p>	Защита лабораторной работы	ОПК-4, ПК-1
<p>Итого по разделу</p>		5	1		9,1			
<p>7. Хранилища данных различных типов. ЦОДы. Витрины данных</p>								

7.1 Хранилища данных (Data Warehouse). Типы хранилищ. ЦОДы. Защищенность, интеграция, агрегация. Разделение наборов данных, используемых для оперативной обработки, и для решения аналитических задач.	3	1	4		1	Самостоятельное изучение учебной и научной литературы, работа с материалами образовательного портала и ЭБС.	Защита лабораторной работы	ПК-1
Итого по разделу		1	4					
8. зачет								
8.1 Подготовка в зачету	3							ОПК-3, ОПК-4, ПК-
Итого по разделу								
Итого за семестр		36	3		34,		зачёт	
Итого по дисциплине		36	6		34, 1		зачет	ОПК-3, ОПК-

## 5 Образовательные технологии

Для реализации предусмотренных видов учебной работы в качестве образовательных технологий в преподавании дисциплины «Информатика» используются традиционная и модульно-компетентностная технологии.

Реализация компетентностного подхода предусматривает использование в учебном процессе активных и интерактивных форм проведения занятий в сочетании с внеаудиторной работой с целью формирования и развития профессиональных навыков обучающихся.

При проведении учебных занятий преподаватель обеспечивает развитие у обучающихся навыков командной работы, межличностной коммуникации, принятия решений, лидерских качеств посредством проведения интерактивных лекций, групповых дискуссий, ролевых игр, тренингов, анализа ситуаций, учета особенностей профессиональной деятельности выпускников и потребностей работодателей.

Формы учебных занятий с использованием традиционных технологий:

- обзорные лекции – для рассмотрения общих вопросов Информатики и информационных технологий, для систематизации и закрепления знаний;
- информационные – для ознакомления с техническими средствами реализации информационных процессов, со стандартами организации сетей, основными приемами защиты информации, и другой справочной информацией;
- лекции-визуализации – для наглядного представления способов решения алгоритмических и функциональных задач, визуализации результатов решения задач; лабораторное занятие, посвященное освоению конкретных умений и навыков по предложенному алгоритму.

Формы учебных занятий с использованием технологий проблемного обучения:

- проблемная - для развития исследовательских навыков и изучения способов решения задач.
- лекции с заранее запланированными ошибками – направленные на поиск студентами синтаксических и алгоритмических ошибок при решении алгоритмических и функциональных задач, с последующей диагностикой слушателей и разбором сделанных ошибок;
- лабораторная работа – организация учебной работы, направленная на решение комплексной учебно-познавательной задачи, требующей от студента применения как научно-теоретических знаний, так и практических навыков;
- занятие на основе кейс-метода – обучение в контексте моделируемой ситуации, воспроизводящей реальные условия научной, производственной, общественной деятельности. Обучающиеся должны проанализировать ситуацию, разобраться в сути проблем, предложить возможные решения и выбрать лучшее из них. Кейсы базируются на реальном фактическом материале или же приближены к реальной ситуации.

Формы учебных занятий с использованием игровых технологий:

- технологии проектного обучения;
- творческий проект – учебно-познавательная деятельность студентов осуществляется в рамках рамочного задания, подчиняясь логике и интересам участников проекта, жанру конечного результата (газета, фильм, праздник, издание, экскурсия, подготовка заданий конкурсов и т.п.);

## **6 Учебно-методическое обеспечение самостоятельной работы обучающихся**

По дисциплине «Основы Data Engineering» предусмотрена аудиторная и внеаудиторная самостоятельная работа обучающихся.

Аудиторная самостоятельная работа обучающихся предполагает выполнение и защиту лабораторных работ, которые определяет для обучающихся преподаватель; выполнения индивидуальных заданий.

Внеаудиторная самостоятельная работа обучающихся осуществляется в виде изучения литературы по соответствующему разделу с проработкой материала; выполнения домашних заданий, подготовки к сдаче лабораторных работ и выполнения домашних заданий с консультациями преподавателя.

**Тема 1.** Основные понятия курса «основы Data Engineering». Технологии обработки Big Data. Отличие от традиционной технологии.

### **Примерные темы для реферата**

1. Понятие Big Data. Роль цифровой информации в 21 веке. Проблемы анализа и обработки большого объема данных.
2. Базовые принципы обработки Big Data.
3. История возникновения термина «Big Data»
4. Модель распределенной обработки данных MapReduce
5. ЦОДы и витрины данных
6. Приемы и стратегии разработки MapReduce-приложений
7. Хранилища данных.
8. Облачные технологии для работы с Big Data.

**Тема 2.** Системы управления Big data.

### **Примерные тестовые вопросы**

1. Объем накопленных человечеством цифровых данных на 2020 год измеряется:
  - Петабайтами
  - Зеттабайтами
  - Экзабайтами
  - Йоттабайтами
2. Перечислите четыре основных характеристики Big Data:
  - Virtualization, Volume, Variability, Vehicle
  - Variety, Velocity, Volume, Value
  - Verification, Volume, Velocity, Visualization
  - Video, Value, Variety, Volume
3. Выберите неверное высказывание:
  - Большие объемы данных приводят к слабой их структуризации, поэтому появляется такое разнообразие данных.
  - Увеличившаяся производительность телекоммуникационных каналов привела к росту объемов передаваемой информации.
  - Удешевление систем хранения на единицу информации привело к росту рынка больших данных.
  - Большое разнообразие источников данных привело к многообразию Big Data.

**Тема 3.** Роль языков программирования Python и R в аналитике больших данных.

### **Примерные вопросы для подготовки к защите лабораторных работ**

1. Перечислите типы данных Python.
2. Вложенные конструкции Python.
3. Процесс анализа. Общая схема анализа. Подготовка данных к анализу.
4. Укажите методы работы со списками и словарями;
5. Работа с текстовыми файлами Python – чтение, запись;
6. Методы работы со строковыми данными;
7. Основы работы с регулярными выражениями;

8. Использование дат в Python: складывание и вычитание даты;
9. Создание и вызов функций; особенности использования;
10. Аналитика Pandas: фильтрование и сортирование данных;
11. Распределение датасет поисковых запросов по количеству слов;

**Задание.** Подготовьте данные к анализу, при необходимости произведите очистку данных.

**Задание.** Используя основные функции библиотеки Pandas провести синтаксический анализ записей файла больших размеров.

**Задание.** С помощью метода merge произведите объединение датафреймов

**Задание.** Произведите группировку данных по колонкам средствами Pandas;

**Задание.** Произведите обогащение выгрузки новыми данными и составьте рейтинг в предметной области.

**Тема 4.** Методы сбора и обработки данных из сети Internet.

#### **Примерные вопросы для подготовки к защите лабораторных работ**

1. Программное обеспечение в области анализа и обработки данных
2. Основные методы сбора информации
3. Технология клиент-сервер
4. Парсинг API
5. Парсинг HTML. BeautifulSoup.
6. Методы класса Data Mining.
7. Кластерный анализ.
8. Краудсорсинг- категоризация и обогащение данных.
9. Смешение и интеграция данных.

**Тема 5.** Организация big data. Реляционные СУБД. СУБД NoSQL. Кластеризация

#### **Примерные вопросы для подготовки к защите лабораторных работ**

1. Аппаратные средства для организации Big Data.
2. Методы для организации Big Data
3. Базы данных SQL и NoSQL-типа.
4. Отличие принципов работы типов СУБД: SQL и NoSQL.
5. Кластеризация. Методы кластеризации.
6. Алгоритмы кластеризации Big Data

**Тема 6.** Инструменты для аналитики Big Data. Системы анализа Big Data

#### **Примерные вопросы для подготовки к защите лабораторных работ**

1. Технологии обработки больших данных: NoSQL, MapReduce, Hadoop.
2. Цели построения и виды распределенных систем
3. Принципы распределенной реализации MapReduce на кластерных системах
4. Функция Map. Функция Reduce. Списки ключ- значение.
5. Технология Hadoop. Применение Hadoop. Файловая система HDFS.
6. Тестирование HDFS.
7. Библиотеки проекта Hadoop.
8. Установка и настройка простого кластера.

**Тема 7.** «Хранилища данных различных типов. ЦОДы. Витрины данных»

#### **Примерные вопросы для подготовки к защите лабораторной работы**

1. Какое оборудование требуется для обработки Больших данных?
2. Центр обработки данных какого уровня обеспечивает максимальную надежность?
3. Центр обработки данных какого уровня обеспечивает резервирование?
4. Центр обработки данных какого уровня позволяет проводить обслуживание оборудования одновременно с обработкой данных?
5. Укажите необходимое время для создания центра обработки данных.
6. Хранилища данных (Data Warehouse). Типы хранилищ. ЦОДы.

## **7 Оценочные средства для проведения промежуточной аттестации**

### ***а) Планируемые результаты обучения и оценочные средства для проведения промежуточной аттестации:***

Промежуточная аттестация по дисциплине включает теоретические вопросы, позволяющие оценить уровень усвоения обучающимися знаний, и практические задания, выявляющие степень сформированности умений и владений; проводится в форме зачета.

Структурный элемент компетенции	Планируемые результаты обучения	Оценочные средства
ОПК-3 способностью применять языки, системы и инструментальные средства программирования в профессиональной деятельности		

Знать	<ul style="list-style-type: none"> <li>- базовые понятия и терминологию Big Data;</li> <li>- возможности инструментальных средств программирования для обработки Big Data;</li> <li>- основные алгоритмы и структуры для работы с Big Data;</li> <li>- основные понятия распределённых систем</li> </ul>	<p style="text-align: center;"><b>Перечень теоретических вопросов:</b></p> <ol style="list-style-type: none"> <li>1. Основные понятия Big Data</li> <li>2. Технологии, используемые для обработки и хранения Big Data.</li> <li>3. Инструментальные системы программирования для обработки больших массивов данных.</li> <li>4. Аналитические процессы обнаружения и обработки ошибок данных.</li> <li>5. Системы управления потоками данных.</li> <li>6. Обработка данных в реальном времени.</li> <li>7. ЦОДы, хранилища данных, витрины данных.</li> <li>8. Теорема CAP и возможность ее обеспечения в реализации распределённых вычислений.</li> <li>9. Масштабируемость баз данных NoSQL</li> <li>10. Основные модели параллельных вычислений.</li> <li>11. Модель вычислений MapReduce.</li> <li>12. Архитектура платформы Apache Hadoop</li> <li>13. Устройство распределенной файловой системы HDFS.</li> <li>14. Технологии поиска Big data: обычный, полнотекстовый и интеграции.</li> <li>15. Параллельные запросы.</li> <li>16. Интеграция данных из различных источников.</li> <li>17. Основные алгоритмы и структуры для работы с Big Data</li> <li>18. Основные функции Python, применяемые для синтаксического анализа Big Data.</li> </ol>
-------	--	---

Уметь	<ul style="list-style-type: none"> <li>- применять инструментальные средства программирования для обработки Big Date;</li> <li>- выявлять скрытые аномалии в данных.</li> </ul>	<p><b>Задание.</b> Используя основные функции библиотеки Pandas провести комбинаторный синтаксический анализ записей файла больших размеров (&gt;30 Гб) вида:</p> <p>X      45.2121  Y      12.7711  Z      -2.2141</p> <p><b>Задание.</b> С помощью алгоритмов и инструментальных средств Питона провести синтаксический анализ записей файла больших размеров (&gt;100 Гб) и произвести подсчет частоты встречаемости слов в наборе данных.</p> <p><b>Задание.</b> Укажите, какие аналитические процессы обнаружения и обработки ошибок данных применимы для указанной структуры документа.</p>
Владеть	<ul style="list-style-type: none"> <li>- современными технологиями обслуживания Big Data;</li> <li>- алгоритмами и методами для обработки Big Date;</li> <li>- инструментальными средствами программирования для обработки Big Date</li> </ul>	<p><b>Задание.</b> Запустите Hadoop кластер из трех нод – 1 главной и 2 подчиненных</p> <p><b>Задание.</b> Создайте программу с реализацией функций Map и Reduce на платформе Apache Hadoop</p> <p><b>Задание.</b> Создайте программу для реализации взаимных блокировок.</p> <p><b>Задание.</b> Имеется набор документов (например, архив новостей). Необходимо для каждого слова, встречающегося в наборе документов, посчитать, сколько раз встречается слово в наборе. Для подсчета используйте технологию MapReduce. Для запуска MapReduce используйте streaming-интерфейс Hadoop. Программы Mapper и Reducer (для выполнения функций Map и Reduce) реализовать на python.</p> <p><b>Задание.</b> Проведите тест по производительности Hadoop MapReduce.</p>
<p>ОПК-4 - способностью понимать значение информации в развитии современного общества, применять достижения современных информационных технологий для поиска информации в компьютерных системах, сетях, библиотечных фондах</p>		

Знать	<ul style="list-style-type: none"> <li>- основные направления развития, значение Big Data и их роль в современном обществе;</li> <li>- основные возможности ИКТ для поиска Big Data из различных источников информации.</li> </ul>	<p><b>Теоретические вопросы:</b></p> <ol style="list-style-type: none"> <li>1. Значение Big Data в развитии современного общества</li> <li>2. Основные направления развития технологий Big Data.</li> <li>3. Современные аппаратные решения для обработки Big Data.</li> <li>4. Основные методы поиска Big Data из различных источников.</li> <li>5. Основные понятия о способах хранения информационных массивов.</li> <li>6. Отличительная особенность NoSQL.</li> <li>7. Назовите основные пять характеристик, присущие Big Data.</li> <li>8. Какие задачи может решать анализ Big Data?</li> <li>9. Назовите принципы разработки параллельных алгоритмов.</li> <li>10. Свойства больших данных и ограничения RDBMS.</li> <li>11. Какие существуют базовые принципы обработки Big Data?</li> <li>12. Дайте определение потоковым данным.</li> <li>13. Что является хранилищем для ключей?</li> </ol>
уметь	<ul style="list-style-type: none"> <li>- организовывать поиск информации в компьютерных системах, сетях и библиотечных фондах.</li> <li>- автоматизировать работу с большими массивами данных, получать данные из внешних источников.</li> </ul>	<p><b>Задание.</b> С помощью информационно-поисковых систем произведите поиск по заданной тематике в структурированных и полуструктурированных данных большого объема (&gt; 5 Гб ) из внешних источников с использованием любой из облачных технологий (Cloud Dataflow, AWS, EMR, Azure). Опишите возможные модели и методы для хранения полученных массивов данных большого объема.</p> <p><b>Задание.</b> Создать программу, которая производит поиск неструктурированных данных в указанной предметной области и создает отчет в соответствии с установленными требованиями.</p>

Владеть	<p>методами поиска, хранения и обработки Big Data из различных источников баз данных, библиотечных фондов;</p> <p>методами представления Big Data с использованием информационных, компьютерных и сетевых технологий</p>	<p><b>Задание.</b> Создайте программу обмена сообщениями между распределенными процессами. В виде распределенной среды выполнения используйте ApacheHadoop.</p> <p><b>Задание.</b> Создайте программу с параллельной сортировкой.</p> <p><b>Задание.</b> Каждый месяц в сети Facebook выкладывается в открытый доступ 30 млрд новых источников информации. Определите, какой необходим объем для хранения полученных данных за год. Используйте технологию MapReduce.</p> <p><b>Задание.</b> Создать программу, которая производит поиск Big Data из Internet сайтов и создает отчет в соответствии с установленными требованиями.</p>
<p>ПК-1 способностью осуществлять поиск, изучение, обобщение и систематизацию научно-технической информации, нормативных и методических материалов в сфере профессиональной деятельности, в том числе на иностранном языке</p>		
Знать	<p>особенности обработки информации в сфере Big Data</p>	<ol style="list-style-type: none"> <li>1. Основные функции библиотеки Pандас для объединения Big Data из различных источников.</li> <li>2. Укажите используемую инфраструктуру для хранения, обработки и передачи Big Data.</li> <li>3. Технологии поиска Big Data: обычный поиск, полнотекстовый и интеграция данных.</li> <li>4. Возможности систематизации Big Data с использованием технологии MapReduce.</li> <li>5. Поиск в структурированных, полуструктурированных и неструктурированных Big Data.</li> <li>6. Технологии облачного хранения Big Data.</li> <li>7. Хранилища данных (Data Warehouse). Типы хранилищ. ЦОДы</li> </ol>
Уметь	<p>обобщать, систематизировать, обеспечивать надежную инфраструктуру полученной научно-технической информации в сфере data инжиниринг.</p>	<p><b>Задача.</b> Представьте витрину данных, которая будет описывать источники данных, используемые учреждением или организацией. Составьте реестр исходных данных с описанием обменных форматов.</p> <p><b>Задача.</b> Опишите метаданные и структуру данных в указанной предметной области. Опишите жизненный цикл метаданных. Проведите анализ контента и оценку требований. Запишите спецификацию системных требований.</p>

Владеть	методами сбора и обобщения научно-технической информации по теоретическим сведениям, проблемам и задачам, решаемым в курсе "Data Инжиниринг", в т.ч. с использованием литературы на иностранном языке.	<p><b>Задача.</b> С помощью информационно-поисковых систем осуществите поиск из внешних источников научно-технической информации в области информационной безопасности на иностранном языке в структуре Hadoop.</p> <p><b>Задание.</b> Определите, какое количество 10 Тб HDD необходимо для хранения набора данных, содержащего координаты, скорости и метаданные (тип объекта и время измерения по указанной характеристике) для всех объектов на территории аэропорта.</p>
---------	--	---

## Критерии оценки для получения зачета

**«зачтено»** – обучающийся показывает средний или высокий уровень сформированности компетенций.

**«не зачтено»** – результат обучения не достигнут, студент не может показать знания на уровне воспроизведения и объяснения информации, не может показать интеллектуальные навыки решения простых задач, не может показать знания на уровне воспроизведения и объяснения информации.

### Вопросы к зачету

1. Значение Big Data в развитии современного общества
2. Основные направления развития технологий Big Data.
3. Современные аппаратные решения для обработки Big Data.
4. Основные методы поиска Big Data из различных источников.
5. Основные понятия о способах хранения информационных массивов.
6. Отличительная особенность NoSQL.
7. Назовите основные пять характеристик, присущие Big Data.
8. Какие задачи может решать анализ Big Data?
9. Назовите принципы разработки параллельных алгоритмов.
10. Свойства больших данных и ограничения RDBMS
11. Какие существуют базовые принципы обработки Big Data?
12. Дайте определение потоковым данным.
13. Что является хранилищем для ключей?
14. Преимущества решений на базе Hadoop.
15. Достоинства и недостатки MapReduce.
16. Распределенные вычисления MapReduce.
17. Распределенные файловые системы.
18. Технологии, используемые для обработки и хранения Big Data.
19. Инструментальные системы программирования для обработки больших массивов данных.
20. Аналитические процессы обнаружения и обработки ошибок данных.
21. Системы управления потоками данных.
22. Обработка данных в реальном времени.
23. ЦОДы, хранилища данных, витрины данных.
24. Масштабируемость баз данных NoSQL
25. Основные модели параллельных вычислений.
26. Модель вычислений MapReduce.
27. Архитектура платформы Apache Hadoop
28. Устройство распределенной файловой системы HDFS.
29. Технологии поиска Big data: обычный, полнотекстовый и интеграции.
30. Параллельные запросы.
31. Интеграция данных из различных источников.
32. Основные алгоритмы и структуры для работы с Big Data

## **8 Учебно-методическое и информационное обеспечение дисциплины**

### **а) Основная литература:**

1. Дадян Э. Г., Зеленков Ю. А. Методы, модели, средства хранения и обработки данных. Учебник для вузов. / Дадян Э. Г., Зеленков Ю. А. Москва: Издательство Вузовский учебник, 2017. - 168 с. ISBN 978-5-9558-0490-3 SBN-онлайн 978-5-16-104262-5 Текст: электронный. - URL: <https://znanium.com/catalog/document?id=53900> (дата обращения: 20.09.2020). – Режим доступа: по подписке.
2. Шелудько, В. М. Язык программирования высокого уровня Python. Функции, структуры данных, дополнительные модули : учебное пособие / В. М. Шелудько ; Южный федеральный университет. - Ростов-наДону ; Таганрог : Издательство Южного федерального университета, 2017. - 107 с. - ISBN 978-5-9275-2648-2. - Текст : электронный. - URL: <https://znanium.com/catalog/product/1021664> (дата обращения: 28.09.2020). – Режим доступа: по подписке.
3. Осипенков, Я. Google Analytics 2019. Полное руководство: практическое руководство / Я. Осипенков. - Москва: ДМК Пресс, 2019. - 748 с. - ISBN 978-5-97060-788-6. - Текст: электронный. - URL: <https://znanium.com/catalog/product/1094922> (дата обращения: 12.10.2020). – Режим доступа: по подписке.
4. Мартишин, С. А. Базы данных: Работа с распределенными базами данных и файловыми системами на примере MongoDB и HDFS с использованием Node.js, Express.js, Apache Spark и Scala : учебное пособие / С. А. Мартишин, В. Л. Симонов, М. В. Храпченко. — Москва : ИНФРА-М, 2020. — 235 с. — (Высшее образование: Бакалавриат). - ISBN 978-5-16-015133-5. - Текст : электронный. - URL: <https://znanium.com/catalog/product/1093646> (дата обращения: 11.10.2020). – Режим доступа: по подписке.

### **б) Дополнительная литература:**

1. Дэвис К., Беликов Д. А. Шаблоны проектирования для облачной среды. Монография. М.: ДМК Пресс, 2020. - 388 с. (высшее образование) ISBN 978-5-97060-807- Текст : электронный. - URL: <https://znanium.com/catalog/document?id=358819> (дата обращения: 20.09.2020). – Режим доступа: по подписке
2. Костюк А. И. Организация облачных и GRID вычислений./ Костюк А. И. Учебное пособие. Издательство: Южный федеральный университет, 2018. - 121 с. (высшее образование) ISBN 978-5-927-52879-0 Текст : электронный. - URL: <https://znanium.com/catalog/document?id=343850> (дата обращения: 20.09.2020). – Режим доступа: по подписке
3. Карпенгер Д., Хьюитт Э., Слинкин А. А. Cassandra. Полное руководство. М.: ДМК Пресс, 2017. - 400 с. ISBN 978-5-97060-453-3 Текст : электронный. - URL:<https://znanium.com/catalog/document?id=341098> (дата обращения: 20.09.2020). – Режим доступа: по подписке
4. Баранкова, И. И. Разработка БД в MS SQL Server с использованием SSMS : учебное пособие / И. И. Баранкова, У. В. Михайлова, Г. И. Лукьянов ; МГТУ. - Магнитогорск : МГТУ, 2018. - 1 электрон. опт. диск (CD-ROM). - Загл. с титул. экрана. - URL: <https://magtu.informsystema.ru/uploader/fileUpload?name=3473.pdf&show=dcatalogues/1/1514290/3473.pdf&view=true> (дата обращения: 04.10.2019). - Макрообъект. - Текст : электронный. - ISBN 978-5-9967-1207-6. - Сведения доступны также на CD-ROM.
5. Маркин, А. В. Программирование на SQL в 2 ч. Часть 1 : учебник и практикум для вузов / А. В. Маркин. — 2-е изд., перераб. и доп. — Москва : Издательство Юрайт, 2020. — 403 с. — (Высшее образование). — ISBN 978-5-534-12256-5. — Текст : электронный // ЭБС Юрайт [сайт]. — URL: <https://urait.ru/bcode/452357> (дата обращения: 20.09.2020).
6. Маркин, А. В. Программирование на SQL в 2 ч. Часть 2 : учебник и практикум для вузов / А. В. Маркин. — 2-е изд., испр. и доп. — Москва : Издательство Юрайт, 2020. — 340 с. — (Высшее образование). — ISBN 978-5-534-12258-9. — Текст : электронный // ЭБС Юрайт [сайт]. — URL: <https://urait.ru/bcode/451185> (дата обращения: 20.09.2020).
7. Лесковец Ю, Раджараман А., Ульман Д., Слинкин А. А. Анализ больших наборов данных. /Лесковец Ю, Раджараман А., Ульман Д., Слинкин А. А. Москва: издательство ДМК Пресс, 2016. - 498 с. (Высшее образование). ISBN 978-5-97060-190-7 Текст: электронный. - URL: <https://znanium.com/catalog/document?id=341090> (дата обращения: 20.09.2020). – Режим доступа: по подписке.
8. Карау Х, Конвински Э., Венделл П., Захария М. Изучаем Spark: молниеносный анализ данных. /Карау Х, Конвински Э., Венделл П., Захария М. Москва: издательство ДМК Пресс, 2015. -314 с. ISBN 978-5-97060-323-9 Текст : электронный. - URL: <https://znanium.com/catalog/document?id=341195> (дата обращения: 20.09.2020). – Режим доступа: по подписке

**в) Методические указания:** см приложение 1

**г) Программное обеспечение и Интернет-ресурсы:****Программное обеспечение**

Наименование ПО	№ договора	Срок действия лицензии
MS Windows 7 Professional(для классов)	Д-1227-18 от 08.10.2018	11.10.2021
Oracle Virtual Box	свободно распространяемое ПО	бессрочно
Oracle SQL Developer Data Modeler	свободно распространяемое ПО	бессрочно
Oracle SQL Developer	свободно распространяемое ПО	бессрочно
Oracle My SQL Workbench Community Edition	свободно распространяемое ПО	бессрочно
MS SQL Server Management Studio	свободно распространяемое ПО	бессрочно
LibreOffice	свободно	бессрочно
Oracle Open	свободно	бессрочно
SCO	свободно	бессрочно
Браузер Mozilla Firefox	свободно распространяемое ПО	бессрочно
Браузер Yandex	свободно	бессрочно
FAR Manager	свободно	бессрочно

**Профессиональные базы данных и информационные справочные системы**

Название курса	Ссылка
Электронная база периодических изданий East View Information Services, ООО	<a href="https://dlib.eastview.com/">https://dlib.eastview.com/</a>
Национальная информационно-аналитическая система – Российский индекс научного цитирования (РИНЦ)	URL: <a href="https://elibrary.ru/project_risc.asp">https://elibrary.ru/project_risc.asp</a>
Поисковая система Академия Google (Google Scholar)	URL: <a href="https://scholar.google.ru/">https://scholar.google.ru/</a>
Информационная система - Единое окно доступа к информационным ресурсам	URL: <a href="http://window.edu.ru/">http://window.edu.ru/</a>
Федеральное государственное бюджетное учреждение «Федеральный институт промышленной собственности»	URL: <a href="http://www1.fips.ru/">http://www1.fips.ru/</a>
Международная реферативная и полнотекстовая справочная база данных	<a href="http://scopus.com">http://scopus.com</a>
Международная база полнотекстовых журналов Springer Journals	<a href="http://link.springer.com/">http://link.springer.com/</a>
Университетская информационная система РОССИЯ	<a href="https://uisrussia.msu.ru">https://uisrussia.msu.ru</a>
Электронные ресурсы библиотеки МГТУ им. Г.И. Носова	<a href="#">sp</a>

Международная база научных материалов в области физических наук и инжиниринга	<a href="http://materials.springer.com/">http://materials.springer.com/</a>
Международная коллекция научных протоколов по различным отраслям знаний	<a href="http://www.springerprotocols.com/">http://www.springerprotocols.com/</a>
Международная база справочных изданий по всем отраслям знаний SpringerReference	<a href="http://www.springer.com/references">http://www.springer.com/references</a>
Международная реферативная и полнотекстовая справочная база данных	<a href="https://www.nature.com/siteindex">https://www.nature.com/siteindex</a>
Архив научных журналов «Национальный электронно-информационный конкорциум» (НП НЭИКОН)	<a href="https://archive.neicon.ru/xmlui/">https://archive.neicon.ru/xmlui/</a>

[www.fstec.ru](http://www.fstec.ru) ; <https://www.rst.gov.ru/portal/gost> ; <https://bdu.fstec.ru/>

### 9 Материально-техническое обеспечение дисциплины (модуля)

Тип и название аудитории	Оснащение аудитории
Учебные аудитории для проведения занятий лекционного типа	Мультимедийные средства хранения, передачи и представления информации.
Учебные аудитории для проведения практических занятий, групповых и индивидуальных консультаций, текущего контроля и промежуточной аттестации	Мультимедийные средства хранения, передачи и представления информации. Комплекс тестовых заданий для проведения промежуточных и рубежных контролей.
Помещения для самостоятельной работы обучающихся	Персональные компьютеры с пакетом MS Office, выходом в Интернет и с доступом в электронную информационно-образовательную среду университета
Помещения для хранения и профилактического обслуживания учебного оборудования	Шкафы для хранения учебно-методической документации, учебного оборудования и учебно-наглядных пособий.

## МЕТОДИЧЕСКИЕ УКАЗАНИЯ ПО ВЫПОЛНЕНИЮ ЛАБОРАТОРНЫХ РАБОТ

Рекомендации направлены на оказание методической помощи студентам при выполнении лабораторных занятий.

**Лабораторное занятие** – это занятие, проводимое под руководством преподавателя в учебной аудитории (компьютерном классе университета или учебной специализированной лаборатории университета), направленное на углубление научно-теоретических знаний и получение лабораторных навыков решения типовых и прикладных задач.

Целью лабораторных занятий является формирование и отработка лабораторных умений и навыков, необходимых в последующей деятельности обучающихся.

Основными задачами лабораторных занятий являются:

- углубление уровня освоения общекультурных и профессиональных компетенций;
- обобщение, систематизация, углубление, закрепление полученных лабораторных знаний по конкретным темам дисциплин различных циклов;
- приобретение студентами умений и навыков использования современных теоретических знаний в решении конкретных прикладных задач;
- развитие профессионального мышления, профессиональной и познавательной мотивации.

Перечень тем лабораторных работ определяется рабочей программой дисциплины. План лабораторных занятий отвечает общей направленности лекционного курса и соотнесен с ним в последовательности тем.

Структура лабораторного занятия включает следующие компоненты: вступительная часть; ответы на вопросы обучающихся; практическая часть; заключительное слово преподавателя. Во вступительной части объявляется тема текущей лабораторной работы, ставится ее цели и задачи, проводится инструктаж по технике безопасности выполнения работы, проверяется исходный уровень готовности студентов к лабораторной работе (выполнение тестов, контрольные вопросы и т.п.), выдается порядок и условия выполнения лабораторной работы.

На лабораторном занятии преподаватель может использовать разнообразные образовательные технологии (методы ИТ, работа в команде, case-study, проблемное обучение, учебные дискуссии и т.п.) по своему выбору для достижения качественного уровня обучения.

### **Правила по технике безопасности для обучающихся при проведении лабораторных работ**

*Общие правила:*

1. Лабораторные работы проводятся под наблюдением преподавателя. К выполнению лабораторных работ студенты допускаются только после прослушивания инструктажа по технике безопасности, правилам поведения, противопожарным мерам в компьютерном классе и специализированных лабораториях.

2. Обучаемый должен строго выполнять правила техники безопасности и санитарно-гигиенические нормы при работе в компьютерных классах и специализированных лабораториях университета.

### **Порядок выполнения лабораторных работ**

При подготовке к выполнению лабораторных работ студент должен повторить теоретический материал, необходимый для выполнения заданий по текущей теме.

Лабораторная работа выполняется каждым студентом самостоятельно, согласно индивидуальному заданию.

Студенты, пропустившие занятия, выполняют лабораторные работы во внеурочное время.

После выполнения каждой лабораторной работы студент демонстрирует результат выполнения преподавателю в виде отчета по лабораторной работе и отвечает на вопросы. Преподаватель оценивает работу в соответствии с заданными критериями оценки лабораторных работ.

### **Правила оформления результатов и оценивания лабораторной работы**

Результаты выполненной лабораторной работы оформляются в соответствии с требованиями к выполнению конкретной работы.

Лабораторная работа считается выполненной, если студент набрал балл, который составляет половину максимального количества баллов.

Для оценивания работы прилагаются следующие критерии.

*Оценка «отлично»* – работа выполнена в полном объеме и без замечаний.

*Оценка «хорошо»* – работа выполнена правильно с учетом 2-3 несущественных ошибок, исправленных самостоятельно по требованию преподавателя.

*Оценка «удовлетворительно»* – работа выполнена правильно не менее чем на половину или допущена существенная ошибка.

*Оценка «неудовлетворительно»* – допущены две (и более) существенные ошибки в ходе работы, которые студент не может исправить по требованию преподавателя, или работа не выполнена.

**МЕТОДИЧЕСКИЕ УКАЗАНИЯ ПО ВЫПОЛНЕНИЮ ВНЕАУДИТОРНЫХ  
САМОСТОЯТЕЛЬНЫХ РАБОТ**

**Общие положения**

Настоящие методические указания предназначены для организации внеаудиторной самостоятельной работы студентов и оказания помощи в самостоятельном изучении теоретического и реализации компетенций обучаемых.

Данные методические указания не являются учебным пособием, поэтому перед началом выполнения самостоятельного задания следует изучить соответствующие разделы лекционных занятий, материалов образовательного портала, разделов основной и дополнительной литературы, представленных в пункте 8. «Учебно-методическое и информационное обеспечение дисциплины (модуля)» данной РПД.

**Цели и задачи самостоятельной работы**

Цель самостоятельной работы – содействие оптимальному усвоению учебного материала обучающимися, развитие их познавательной активности, готовности и потребности в самообразовании.

**Задачи самостоятельной работы:**

- повышение исходного уровня владения информационными технологиями;
- углубление и систематизация знаний;
- постановка и решение стандартных задач профессиональной деятельности;
- развитие работы с различной по объему и виду информацией, учебной и научной литературой;
- практическое применение знаний, умений;
- самостоятельно использование стандартных программных средств сбора, обработки, хранения и защиты информации
- развитие навыков организации самостоятельного учебного труда и контроля за его эффективностью.

Виды внеаудиторной самостоятельной работы и формы контроля и время на выполнение каждого вида самостоятельной работы указаны в пункте 4. «Структура и содержание дисциплины» данной РПД.

**Порядок выполнения**

При выполнении текущей внеаудиторной самостоятельной работы обучающемуся следует придерживаться следующего порядка действий:

- 1) внимательно изучить соответствующие теоретические разделы дисциплины, пользуясь материалами (лекционными, презентационными, аудио-визуальными):
  - а) предоставляемыми преподавателем на лекционных занятиях;
  - б) предоставляемыми преподавателем в рамках электронных образовательных курсов;
  - в) содержащимися в учебниках и учебных пособиях ЭБС (электронно-библиотечных систем), электронных каталогов университета и интернет-ресурсов.
- 2) Подробно разобрать типовые примеры решения задач, рассмотренные в рамках аудиторной контактной работы с преподавателем.
- 3) Применить полученные теоретические знания и практические навыки к решению индивидуальных заданий, к прохождению компьютерных тестирований.
- 4) При необходимости, сформировать перечень вопросов, вызвавших затруднения в процессе самостоятельной работы. Обсудить возникшие вопросы со студентами группы, в рамках командно-проектной работы, и с преподавателем, в рамках консультационной помощи, реализованной либо в контактной форме, либо средствами информационно-образовательной среды ВУЗа.

## **Критерии оценки внеаудиторных самостоятельных работ**

Качество выполнения внеаудиторной самостоятельной работы обучающихся оценивается посредством текущего контроля самостоятельной работы обучающихся с использованием балльно-рейтинговой системы.

В качестве форм текущего контроля по дисциплине используются: индивидуальные задания, аудиторские контрольные работы, компьютерное тестирование.

Максимальное количество баллов обучающийся получает, если:

- выполняет индивидуальные задания в соответствии со всеми заявленными требованиями;
- дает правильные формулировки, точные определения, понятия терминов;
- может обосновать рациональность решения текущей задачи.;
- обстоятельно с достаточной полнотой излагает соответствующую теоретический раздел;
- правильно отвечает на дополнительные вопросы преподавателя, имеющие целью выяснить степень понимания им данного материала.

50~85% от максимального количества баллов обучающийся получает, если:

- неполно (не менее 70% от полного), но правильно выполнено задание;
- при изложении были допущены 1-2 несущественные ошибки, которые он исправляет после замечания преподавателя;
- дает правильные формулировки, точные определения, понятия терминов;
- может обосновать свой ответ, привести необходимые примеры;
- правильно отвечает на дополнительные вопросы преподавателя, имеющие целью выяснить степень понимания им данного материала.

36~50% от максимального количества баллов обучающийся получает, если:

- неполно (не менее 50% от полного), но правильно изложено задание;
- при изложении была допущена 1 существенная ошибка;
- знает и понимает основные положения данной темы, но допускает неточности в формулировке понятий;
- излагает выполнение задания недостаточно логично и последовательно;
- затрудняется при ответах на вопросы преподавателя.

35% и менее от максимального количества баллов обучающийся получает, если:

- неполно (менее 50% от полного) изложено задание;
- при изложении были допущены существенные ошибки. В "0" баллов преподаватель вправе оценить выполненное обучающимся задание, если оно не удовлетворяет требованиям, установленным преподавателем к данному виду работы или не было представлено для проверки.

Сумма полученных баллов по всем видам заданий внеаудиторной самостоятельной работы составляет рейтинговый показатель обучающегося. Рейтинговый показатель обучающегося влияет на выставление итоговой оценки по результатам изучения дисциплины.

Показатели и критерии оценивания полученных знаний представлены в пункте 7.6) «Оценочные средства для проведения промежуточной аттестации» данной РПД.

**Основная литература:** бакалавриат, но в аннотации есть и магистратура, и аспирантура



## Аннотация

В учебнике рассмотрены следующие вопросы: теория, методология и практика работы с данными, базами данных, базами знаний и хранилищами данных; модели представления данных в памяти компьютера — модель «сущность — связь»; дореляционные модели представления данных, реляционные базы данных; проектирование реляционных баз данных; направления развития баз данных; физическая организация систем управления базами данных. Сформулированы концептуальные основы понятия «знание», изложены понятия и определения знаний, баз знаний, модели представления знаний. Здесь же даны принципы построения систем, ориентированных на анализ данных, — хранилища данных, модели данных, используемые при построении хранилищ данных. Рассмотрены основные вопросы защиты данных, обеспечения их безопасности и секретности, а также системы хранения и обработки больших данных. Учебник предназначен для бакалавров и магистрантов всех специальностей, а также для аспирантов и слушателей ИПК.

Шелудько, В. М. Язык программирования высокого уровня Python. Функции, структуры данных, дополнительные модули : учебное пособие / В. М. Шелудько ; Южный федеральный университет. - Ростов-наДону ; Таганрог : Издательство Южного федерального университета, 2017. - 107 с. - ISBN 978-5-9275-2648-2. - Текст : электронный. - URL: <https://znanium.com/catalog/product/1021664> (дата обращения: 28.09.2020). – Режим доступа: по подписке.

znanium.com электронно-библиотечная система

[Каталог](#) [Коллекции](#) [О проекте](#) [Цены](#) [Помощь](#) [Регистрация](#) [Вход](#)

Главная » [Каталог](#) » [Книги](#) » Язык программирования высокого уровня Python. Функции, структуры данных, дополнительные модули


[Расширенный поиск](#)

Новинки

Книги ▼

Журналы ▼

Статьи ▼

Авторы

Издательства

Аффилиация авторов

Подборки Znanium

Свободный доступ

Публичным библиотекам

Книги с доп. материалами



**Язык программирования высокого уровня Python. Функции, структуры данных, дополнительные модули**

Нет доступа Основная коллекция

Издательство: Южный федеральный университет

Вид издания: Учебное пособие

Уровень образования: ВО - Специалитет

Авторы: Шелудько Виктория Михайловна

Год издания: 2017

Кол-во страниц: 107

ISBN: 978-5-9275-2648-2

Артикул: 708971.01.99

[Ознакомиться](#)

Доступ онлайн

**128 руб.**

В корзину

**Аннотация**

В учебном пособии рассматриваются принципы разработки пользовательских процедур, структур данных и модулей на языке высокого уровня Python. Дается краткое описание основных модулей языка. Большое количество наглядных примеров позволит освоить основные принципы составления программы на языке Python.

znanium.com электронно-библиотечная система

[Каталог](#) [Коллекции](#) [О проекте](#) [Цены](#) [Помощь](#) [Регистрация](#) [Вход](#)

Главная » [Каталог](#) » [Книги](#) » Базы данных. Практическое применение СУБД SQL и NoSQL-типа для применения проектирования информационных систем


[Расширенный поиск](#)

Новинки

Книги ▼

Журналы ▼

Статьи ▼

Авторы

Издательства

Аффилиация авторов

Подборки Znanium

Свободный доступ

Публичным библиотекам

Книги с доп. материалами



**Базы данных. Практическое применение СУБД SQL и NoSQL-типа для применения проектирования информационных систем**

Нет доступа Основная коллекция

Издательство: ФОРУМ

Вид издания: Учебное пособие

Уровень образования: ВО - Бакалавриат

Авторы: Мартишин Сергей Анатольевич, Симонов Владимир Львович, Храпченко Марина Валерьевна

Год издания: 2020

Кол-во страниц: 368

ISBN: 978-5-8199-0718-4

ISBN-онлайн: 978-5-16-104936-5

Артикул: 633184.06.01

[Ознакомиться](#)

Доступ онлайн

от **444 руб.**

В корзину

Бумажная книга

**1 512 руб.**

Купить

**Аннотация**

Предназначено для изучения и создания информационных систем и приложений с использованием распределенных баз данных и файловых систем. Для создания информационных систем на основе распределенных баз данных используется серверная часть стека программного обеспечения MEAN, включающая в себя СУБД NoSQL-типа MongoDB, платформу Node и фреймворк Express. Рассмотрены также основы администрирования и репликации в СУБД MongoDB. В качестве инструмента моделирования объектов баз данных используется Mongoose.js. С целью создания фронтенд частей сайтов и веб-приложений изучается применение свободного программного обеспечения Bootstrap, HTML5 Boilerplate, Initializr. Приводится детальный процесс разработки и тестирования API REST-приложения на основе модели MVC. В результате обучающиеся будут способны создавать полноценную распределенную информационную систему на основе современных СУБД NoSQL-типа. Для создания приложений также используются распределенные хранилища данных (больших массивов неструктурированной информации) на примере Hadoop Apache с инструментарием Spark и Scala. **Предназначено для бакалавров, магистрантов, аспирантов** и специалистов в области проектирования распределенных информационных систем на основе современных СУБД NoSQL-типа, создания мобильных и интернет-приложений. Также полезно для лиц, занимающихся самообразованием.

Осипенков, Я. Google Analytics 2019. Полное руководство : практическое руководство / Я. Осипенков. - Москва : ДМК Пресс, 2019. - 748 с. - ISBN 978-5-97060-788-6. - Текст : электронный. - URL: <https://znanium.com/catalog/product/1094922> (дата обращения: 28.09.2020). – Режим доступа: по подписке.

### Аннотация

В данной книге вы найдете исчерпывающие материалы по настройке и работе с самым популярным инструментом веб-аналитики в мире. Подробно рассматриваются все тонкости работы с продуктом, описываются пошаговые алгоритмы настройки, разбираются основные функции и интерфейс, даются рекомендации по построению отчетов с практическими примерами их использования. Издание будет полезно всем, кто хочет изучить Google Analytics самостоятельно, быстро погрузиться в процесс анализа данных, а также получить базовые знания в построении и автоматизации отчетности.

The screenshot shows the Znanium.com website interface. At the top, there is a navigation bar with the site logo and links for 'Каталог', 'Коллекции', 'О проекте', 'Цены', and 'Помощь'. Below the navigation bar, there is a search bar and a breadcrumb trail: 'Главная > Каталог > Книги > Google Analytics 2019. Полное руководство'. The main content area is divided into two columns. The left column contains a vertical menu with categories: 'Новинки', 'Книги', 'Журналы', 'Статьи', 'Авторы', 'Издательства', 'Аффилиация авторов', 'Подборки Znanium', 'Свободный доступ', 'Публичным библиотекам', and 'Книги с доп. материалами'. The right column displays the product details for 'Google Analytics 2019. Полное руководство'. It includes a book cover image, a 'Нет доступа' (No access) status, and a 'Новинка' (New) badge. The publisher is 'ДМК Пресс'. The book is a 'Практическое пособие' (Practical manual) for 'ВО - Специалист' (Higher education - Specialist) level. The author is 'Осипенков Яков Михайлович'. The edition year is 2019, with 748 pages. The ISBN is 978-5-97060-788-6, and the article number is 739781.01.99. Below the product details, there is an 'Аннотация' (Annotation) section with the same text as provided in the input.

This screenshot shows a web browser window displaying the Znanium.com website. At the top, there is a notification banner from Magnitogorsk State Technical University: 'Вы находитесь в сети организации Магнитогорский государственный технический университет им. Г.И. Носова. Для получения возможности удаленной работы с сайтом, пожалуйста, авторизуйтесь.' Below the notification, the website's navigation and search elements are visible. The main content area shows a product page for 'Анализ больших наборов данных' (Analysis of large data sets). The product is marked as 'Нет доступа' (No access) and has an online price of 4,000 rubles. The publisher is 'ДМК Пресс'. The book is a 'Практическое пособие' (Practical manual) for 'ВО - Бакалавриат' (Higher education - Bachelor's) level. The authors are 'Лесковец Юре, Раджараман Ананд, Ульман Джеффри Д., Слинкин А. А.'. The edition year is 2016, with 498 pages. The ISBN is 978-5-97060-190-7, and the article number is 688462.02.99. The browser's taskbar at the bottom shows various open applications and the system clock indicating 21:11 on 30.09.2020.

Эта книга написана ведущими специалистами в области технологий баз данных и веба. Благодаря популярности интернет-торговли появилось много чрезвычайно объемных баз данных,

для извлечения информации из которых нужно применять методы добычи данных (data mining). В книге описываются алгоритмы, которые реально использовались для решения важнейших задач добычи данных и могут быть с успехом применены даже к очень большим наборам данных. Изложение начинается с рассмотрения технологии MapReduce - важного средства распараллеливания алгоритмов. Излагаются алгоритмы хэширования с учетом близости и потоковой обработки данных, которые поступают слишком быстро для тщательного анализа. В последующих главах рассматривается идея показателя PageRank, нахождение частых предметных наборов и кластеризация. Во второе издание включен дополнительный материал о социальных сетях, машинном обучении и понижении размерности. Издание будет в равной мере полезна студентам и программистам-практикам.

Мартишин, С. А. Базы данных: Работа с распределенными базами данных и файловыми системами на примере MongoDB и HDFS с использованием Node.js, Express.js, Apache Spark и Scala : учебное пособие / С. А. Мартишин, В. Л. Симонов, М. В. Храпченко. — Москва : ИНФРА-М, 2020. — 235 с. — (Высшее образование: **Бакалавриат**). - ISBN 978-5-16-015133-5. - Текст : электронный. - URL: <https://znanium.com/catalog/product/1093646> (дата обращения: 11.10.2020). — Режим доступа: по подписке.

### В Аннотации- магистратура и аспирантура

Предназначено для изучения и создания информационных систем и приложений с использованием распределенных баз данных и файловых систем. Для создания информационных систем на основе распределенных баз данных используется серверная часть стека программного обеспечения MEAN, включающая в себя СУБД NoSQL-типа MongoDB, платформу Node и фреймворк Express. Рассмотрены также основы администрирования и репликации в СУБД MongoDB. В качестве инструмента моделирования объектов баз данных используется Mongoose.js. С целью создания фронтенд частей сайтов и веб-приложений изучается применение свободного программного обеспечения Bootstrap, HTML5 Boilerplate, Initializr. Приводится детальный процесс разработки и тестирования API REST-приложения на основе модели MVC. В результате обучающиеся будут способны создавать полноценную распределенную информационную систему на основе современных СУБД NoSQL-типа. Для создания приложений также используются распределенные хранилища данных (больших массивов неструктурированной информации) на примере Hadoop Apache с инструментарием Spark и Scala. Предназначено **для бакалавров, магистрантов, аспирантов** и специалистов в области проектирования распределенных информационных систем на основе современных СУБД NoSQL-типа, создания мобильных и интернет-приложений. Также полезно для лиц, занимающихся самообразованием.

<https://znanium.com/catalog/document?id=350398>

znanium.com/catalog/document?id=350398

znanium.com электронно-библиотечная система

Каталог Коллекции О проекте Цены Помощь Регистрация Вход

Главная > Каталог > Книги > Базы данных. Практическое применение СУБД SQL и NoSQL-типа для применения проектирования информационных систем

Поиск изданий в каталоге Расширенный поиск

Новинки Книги Журналы Статьи Авторы Издательства Афилиация авторов Подборки Znanium Свободный доступ Публичным библиотекам Книги с доп. материалами



**Базы данных. Практическое применение СУБД SQL и NoSQL-типа для применения проектирования информационных систем**

Доступ онлайн от 444 руб

Нет доступа Основная коллекция

В корзину

Бумажная книга 1512 руб

Купить

Издательство: ФОРУМ  
Вид издания: Учебное пособие  
Уровень образования: ВО - Бакалавриат  
Авторы: Мартишин Сергей Анатольевич, Симонов Владимир Львович, Храпченко Марина Валерьевна

Год издания: 2020  
Кол-во страниц: 368  
ISBN: 978-5-8199-0718-4  
ISBN-онлайн: 978-5-16-104936-5  
Артикул: 633184.06.01

Ознакомиться

Аннотация

Учебное пособие предназначено для изучения практического применения СУБД SQL и NoSQL-типа при проектировании информационных систем. Проведены аналогии между базами данных SQL и NoSQL-типа. Полученные теоретические знания закрепляются при выполнении цикла лабораторных работ. Работа с SQL СУБД изучается на примере СУБД MariaDB, рассмотрен вопрос использования хорошо зарекомендовавших себя в работе с MySQL приложений (например, phpMyAdmin, MySQL Workbench). Работа с NoSQL СУБД изучается на примере СУБД MongoDB с использованием оболочки Robotomongo. Показана возможность использования универсального менеджера баз данных DBeaver для одновременной работы с СУБД MariaDB и MongoDB. Приведены примеры информационных систем на базе обоих типов СУБД SQL и NoSQL. Предназначено для студентов, бакалавров, **магистрантов, аспирантов** и специалистов в области проектирования информационных систем и баз данных. Пособие также полезно для лиц, занимающихся самообразованием.