

Министерство науки и высшего образования Российской Федерации  
Федеральное государственное автономное образовательное учреждение  
высшего образования  
«Уральский федеральный университет имени первого Президента России Б.Н. Ельцина»

УТВЕРЖДАЮ  
Директор по образовательной  
деятельности

\_\_\_\_\_ С.Т. Князев  
«\_\_» \_\_\_\_\_

### РАБОЧАЯ ПРОГРАММА МОДУЛЯ

<b>Код модуля</b>	<b>Модуль</b>
1153907	Базы данных

Екатеринбург

<b>Перечень сведений о рабочей программе модуля</b>	<b>Учетные данные</b>
<b>Образовательная программа</b> 1. Прикладной анализ данных	<b>Код ОП</b> 1. 09.04.02/33.01
<b>Направление подготовки</b> 1. Информационные системы и технологии	<b>Код направления и уровня подготовки</b> 1. 09.04.02

Программа модуля составлена авторами:

<b>№ п/п</b>	<b>Фамилия Имя Отчество</b>	<b>Ученая степень, ученое звание</b>	<b>Должность</b>	<b>Подразделение</b>
1	Бурак Илья Александрович	без ученой степени, без ученого звания	Преподаватель	ПАО "Сбербанк"
2	Корелин Иван Андреевич	кандидат технических наук	Доцент	Базовая кафедра "Автоматизация финансовых систем"
3	Медведев Александр Николаевич	к.т.н., с.н.с.	доцент	Базовая кафедра "Аналитика больших данных и методы видеоанализа"
4	Созыкин Андрей Владимирович	кандидат технических наук, без ученого звания	Доцент	информационных технологий и систем управления

**Согласовано:**

Управление образовательных программ

Р.Х. Токарева

# 1. ОБЩАЯ ХАРАКТЕРИСТИКА МОДУЛЯ Базы данных

## 1.1. Аннотация содержания модуля

Модуль «Базы данных» состоит из трех дисциплин – «Базы данных», «Инструменты Business Intelligence» и «Обработка больших объемов данных». Содержание дисциплин модуля позволит студентам изучить комплексы инструментов бизнес-аналитики для обработки, преобразования, анализа и визуализации данных. Изучение дисциплин модуля позволит студентам овладеть необходимыми знаниями и умениями для практического освоения методов создания и разработки систем управления базами данных.

## 1.2. Структура и объем модуля

Таблица 1

№ п/п	Перечень дисциплин модуля в последовательности их освоения	Объем дисциплин модуля и всего модуля в зачетных единицах
1	Базы данных	6
2	Инструменты Business Intelligence	6
3	Обработка больших объемов данных	6
ИТОГО по модулю:		18

## 1.3. Последовательность освоения модуля в образовательной программе

Пререквизиты модуля	Не предусмотрены
Постреквизиты и кореквизиты модуля	Не предусмотрены

## 1.4. Распределение компетенций по дисциплинам модуля, планируемые результаты обучения (индикаторы) по модулю

Таблица 2

Перечень дисциплин модуля	Код и наименование компетенции	Планируемые результаты обучения (индикаторы)
1	2	3
Базы данных	ПК-4 - Способен осуществлять управление развитием баз данных	3-1 - Привести примеры основных тенденций развития информационных технологий в области баз данных (БД)

		<p>З-3 - Сформулировать отличительные особенности современных и перспективных БД</p> <p>У-1 - Прогнозировать состояние и осуществлять планирование по развитию БД</p> <p>У-3 - Планировать мероприятия по переходу на новую версию БД</p> <p>П-2 - Разрабатывать рекомендации по освоению и внедрению в практику администрирования новых технологий работы с БД</p>
	ПК-7 - Способен разрабатывать системы управления базами данных	<p>З-1 - Сделать обзор методов тестирования современных систем управления базами данных</p> <p>З-4 - Понимать специальную терминологию в области систем управления базами данных</p> <p>У-3 - Анализировать и применять нормативно-техническую документацию при использовании систем управления базами данных</p> <p>П-1 - Осуществлять обоснованный анализ ошибок в компонентах системы управления базами данных, в соответствии с данными эксплуатации</p> <p>П-3 - Иметь практический опыт разработки системы управления базами данных</p>
Инструменты Business Intelligence	ПК-4 - Способен осуществлять управление развитием баз данных	<p>З-2 - Сделать обзор современных и перспективных технологий в области БД</p> <p>З-3 - Сформулировать отличительные особенности современных и перспективных БД</p> <p>У-2 - Анализировать и соотносить возможности внедрения новых информационных технологий в проект</p> <p>П-1 - Осуществить исследование рынка перспективных БД и подготовить презентацию их возможностей</p>
	ПК-7 - Способен разрабатывать системы управления базами данных	<p>З-3 - Характеризовать методы обработки данных</p> <p>З-4 - Понимать специальную терминологию в области систем управления базами данных</p>

		<p>У-3 - Анализировать и применять нормативно-техническую документацию при использовании систем управления базами данных</p> <p>П-3 - Иметь практический опыт разработки системы управления базами данных</p>
Обработка больших объемов данных	<p>ПК-1 - Способен использовать методы и средства системной инженерии в области получения, передачи, хранения, переработки и представления информации посредством информационных технологий</p>	<p>З-1 - Изложить основные методы и средства системной инженерии в области получения, передачи, хранения, переработки и представления информации</p> <p>З-2 - Характеризовать особенности применения информационных технологий в области получения, передачи, хранения, переработки и представления информации</p> <p>У-1 - Выбирать адекватные методы и средства системной инженерии в области получения, передачи, хранения, переработки и представления информации</p> <p>У-2 - Систематизировать и оценивать научно-техническую информацию о методах цифровой обработки сигналов для переработки и представления информации</p> <p>П-1 - Осуществлять обоснованный выбор методов цифровой обработки сигналов для переработки и представления информации посредством информационных технологий с учетом особенностей сигналов и изображений</p>
	<p>ПК-7 - Способен разрабатывать системы управления базами данных</p>	<p>З-2 - Классифицировать основные структуры данных</p> <p>З-3 - Характеризовать методы обработки данных</p> <p>У-2 - Определять оптимальные методы обработки данных</p> <p>П-2 - Сделать вывод об оптимальности применения методов обработки данных для системы управления базами данных</p>

### 1.5. Форма обучения

Обучение по дисциплинам модуля может осуществляться в очной формах.

**РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ**  
**Базы данных**

Рабочая программа дисциплины составлена авторами:

<b>№ п/п</b>	<b>Фамилия Имя Отчество</b>	<b>Ученая степень, ученое звание</b>	<b>Должность</b>	<b>Подразделение</b>
1	Бурак Илья Александрович	без ученой степени, без ученого звания	Преподаватель	ПАО "Сбербанк"
2	Созыкин Андрей Владимирович	кандидат технических наук, без ученого звания	Доцент	информационных технологий и систем управления

**Рекомендовано учебно-методическим советом института** Радиоэлектроники и информационных технологий - РТФ

Протокол № 4 от 28.05.2021 г.

# 1. СОДЕРЖАНИЕ И ОСОБЕННОСТИ РЕАЛИЗАЦИИ ДИСЦИПЛИНЫ

Авторы:

- Бурак Илья Александрович, Преподаватель, ПАО "Сбербанк"
- Созыкин Андрей Владимирович, Доцент, информационных технологий и систем управления

## 1.1. Технологии реализации, используемые при изучении дисциплины модуля

- Традиционная (репродуктивная) технология
- Разноуровневое (дифференцированное) обучение
  - Базовый уровень

*\*Базовый I уровень – сохраняет логику самой науки и позволяет получить упрощенное, но верное и полное представление о предмете дисциплины, требует знание системы понятий, умение решать проблемные ситуации. Освоение данного уровня результатов обучения должно обеспечить формирование запланированных компетенций и позволит обучающемуся на минимальном уровне самостоятельности и ответственности выполнять задания;*  
*Продвинутый II уровень – углубляет и обогащает базовый уровень как по содержанию, так и по глубине проработки материала дисциплины. Это происходит за счет включения дополнительной информации. Данный уровень требует умения решать проблемы в рамках курса и смежных курсов посредством самостоятельной постановки цели и выбора программы действий. Освоение данного уровня результатов обучения позволит обучающемуся повысить уровень самостоятельности и ответственности до творческого применения знаний и умений.*

## 1.2. Содержание дисциплины

Таблица 1.1

Код раздела, темы	Раздел, тема дисциплины*	Содержание
1	Основные модели данных	Типы данных в БД. Иерархическая модель БД. Сетевая модель БД. Общие понятия БД.
2	Реляционные БД	Теория нормальных форм. Диаграммы представления данных. Проектирование логической структуры БД.
3	Язык SQL	Язык SQL. Работа со строками. Агрегация данных. Многотабличные запросы. SQL. CRUD
4	СУБД Oracle	Особенности структуры СУБД Oracle. Индексация. План запросов.
5	Язык PL/SQL	PL/SQL. Расширение. PL/SQL. Анонимные блоки. PL/SQL. Процедуры, функции, пакеты.

## 1.3. Направление, виды воспитательной деятельности и используемые технологии

Направления воспитательной деятельности сопрягаются со всеми результатами обучения компетенций по образовательной программе, их освоение обеспечивается содержанием всех дисциплин модулей.

1.4. Программа дисциплины реализуется на государственном языке Российской Федерации .

## **2. УЧЕБНО-МЕТОДИЧЕСКОЕ И ИНФОРМАЦИОННОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ**

### **Базы данных**

#### **Электронные ресурсы (издания)**

1. Аврунев, О. Е.; Модели баз данных : учебное пособие.; Новосибирский государственный технический университет, Новосибирск; 2018; <https://biblioclub.ru/index.php?page=book&id=575324> (Электронное издание)
2. Зудилова, , Т. В.; SQL и PL/SQL для разработчиков СУБД Oracle; Университет ИТМО, Санкт-Петербург; 2012; <http://www.iprbookshop.ru/65745.html> (Электронное издание)
3. Кузнецов, , С. Д.; Введение в модель данных SQL : учебное пособие.; Интернет-Университет Информационных Технологий (ИНТУИТ), Ай Пи Ар Медиа, Москва; 2021; <http://www.iprbookshop.ru/101995.html> (Электронное издание)
4. Пржиялковский, , В. В.; Введение в Oracle SQL : учебное пособие.; Интернет-Университет Информационных Технологий (ИНТУИТ), Ай Пи Ар Медиа, Москва; 2020; <http://www.iprbookshop.ru/94846.html> (Электронное издание)

#### **Профессиональные базы данных, информационно-справочные системы**

1. Applied Science & Technology Source. EBSCO publishing <http://search.ebscohost.com>
2. Wiley Online Library <http://onlinelibrary.wiley.com/>
3. Гугл Академия <https://scholar.google.ru/>

#### **Материалы для лиц с ОВЗ**

Весь контент ЭБС представлен в виде файлов специального формата для воспроизведения синтезатором речи, а также в тестовом виде, пригодном для прочтения с использованием экранной лупы и настройкой контрастности.

#### **Базы данных, информационно-справочные и поисковые системы**

1. Academic Search Ultimate EBSCO publishing – <http://search.ebscohost.com>
2. eBook Collections Springer Nature – <https://link.springer.com/>
3. Гугл Академия – <https://scholar.google.ru/>
4. Электронный научный архив УрФУ <https://elar.urfu.ru/>
5. Зональная научная библиотека (УрФУ) - <http://lib.urfu.ru/>
6. Портал информационно-образовательных ресурсов УрФУ <https://study.urfu.ru/>



7. Электронно-библиотечная система «Лань» – <https://e.lanbook.com/>
8. Университетская библиотека ONLINE – <https://biblioclub.ru/>
9. Электронно-библиотечная система "Библиокомплектатор" (IPRbooks)  
<http://www.bibliocomplectator.ru/available>
10. Электронные информационные ресурсы Российской государственной библиотеки  
<https://www.rsl.ru/>
11. Научная электронная библиотека «КиберЛенинка» <https://cyberleninka.ru/>

### 3. МАТЕРИАЛЬНО-ТЕХНИЧЕСКОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ

#### Базы данных

#### Сведения об оснащённости дисциплины специализированным и лабораторным оборудованием и программным обеспечением

Таблица 3.1

№ п/п	Виды занятий	Оснащённость специальных помещений и помещений для самостоятельной работы	Перечень лицензионного программного обеспечения
1	Лекции	Мебель аудиторная с количеством рабочих мест в соответствии с количеством студентов Рабочее место преподавателя Доска аудиторная Оборудование, соответствующее требованиям организации учебного процесса в соответствии с санитарными правилами и нормами Подключение к сети Интернет	Adobe Acrobat Professional 2017 Multiple Platforms Office 365 EDUA3 ShrdSvr ALNG SubsVL MVL PerUsr B Faculty EES В курсе используется бесплатно распространяемое программное обеспечение: Oracle Database XE, Oracle SQL Developer.
2	Практические занятия	Мебель аудиторная с количеством рабочих мест в соответствии с количеством студентов Рабочее место преподавателя Доска аудиторная Периферийное устройство Персональные компьютеры по количеству обучающихся Оборудование, соответствующее требованиям	Office 365 EDUA3 ShrdSvr ALNG SubsVL MVL PerUsr B Faculty EES В курсе используется бесплатно распространяемое программное обеспечение: Oracle Database XE, Oracle SQL Developer.

		организации учебного процесса в соответствии с санитарными правилами и нормами Подключение к сети Интернет	
--	--	---	--

**РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ**  
**Инструменты Business Intelligence**

Рабочая программа дисциплины составлена авторами:

<b>№ п/п</b>	<b>Фамилия Имя Отчество</b>	<b>Ученая степень, ученое звание</b>	<b>Должность</b>	<b>Подразделение</b>
1	Медведев Александр Николаевич	к.т.н., с.н.с.	доцент	Базовая кафедра "Аналитика больших данных и методы видеоанализа"
2	Созыкин Андрей Владимирович	кандидат технических наук, без ученого звания	Доцент	информационных технологий и систем управления

**Рекомендовано учебно-методическим советом института Радиозлектроники и информационных технологий - РТФ**

Протокол № 4 от 28.05.2021 г.

# 1. СОДЕРЖАНИЕ И ОСОБЕННОСТИ РЕАЛИЗАЦИИ ДИСЦИПЛИНЫ

Авторы:

- **Медведев Александр Николаевич, доцент, Базовая кафедра "Аналитика больших данных и методы видеоанализа"**
- **Созыкин Андрей Владимирович, Доцент, информационных технологий и систем управления**

## 1.1. Технологии реализации, используемые при изучении дисциплины модуля

- Традиционная (репродуктивная) технология
- Разноуровневое (дифференцированное) обучение
  - Базовый уровень

*\*Базовый I уровень – сохраняет логику самой науки и позволяет получить упрощенное, но верное и полное представление о предмете дисциплины, требует знание системы понятий, умение решать проблемные ситуации. Освоение данного уровня результатов обучения должно обеспечить формирование запланированных компетенций и позволит обучающемуся на минимальном уровне самостоятельности и ответственности выполнять задания;*

*Продвинутый II уровень – углубляет и обогащает базовый уровень как по содержанию, так и по глубине проработки материала дисциплины. Это происходит за счет включения дополнительной информации. Данный уровень требует умения решать проблемы в рамках курса и смежных курсов посредством самостоятельной постановки цели и выбора программы действий. Освоение данного уровня результатов обучения позволит обучающемуся повысить уровень самостоятельности и ответственности до творческого применения знаний и умений.*

## 1.2. Содержание дисциплины

Таблица 1.1

Код раздела, темы	Раздел, тема дисциплины*	Содержание
P1	Введение в область бизнес-аналитики	Обзор областей применения бизнес аналитики, х особенности и преимущества использования BI систем как для руководящих позиций, так и для рядовых аналитиков. Обзор и анализ инструментов бизнес-анализа
P2	Базовые инструменты бизнес- анализа	Введение в работу с Power BI. Сравнение возможностей Power BI и Excel. Введение в работу с Tableau. Сравнение возможностей Power BI и Tableau.
P3	Методика публикаций работ бизнес аналитика	Изучение способов публикации дашбордов в Power BI и Tableau.
P4	Информационные ресурсы для публикаций аналитических исследований	Работа с галереями Microsoft Power BI Gallery. Работа с галереями Tableau Gallery
P5	Методики проведения исследований и опросов	Разбор исследовательских дашбордов и принципов составления хорошей формы сбора информации.
P6	Интеллектуальная обработка данных и выделение	Анализ полученной неструктурированной информации методами Power BI, Excel и Google Sheets

	наиболее важной информации	
<b>P7</b>	Метод презентации аналитических дашбордов широкой аудитории	Разбор методик презентаций проектов AIDA и Elevator Speech. Практика в представлении проекта с помощью методики AIDA.
<b>P8</b>	Обзор продуктов и технологических решений SAP	Введение в аналитические решения SAP на базе SAP HANA (Введение в анализ данных в SAP HANA). Machine Learning in SAP HANA. Моделирование и анализ данных. SAP HANA Data Modeling Основы Data Mining и предиктивная аналитика SAP Predictive Analytics Работа со встроенными алгоритмами PAL (Predictive Analysis Library) Введение в обработку графов (Introduction to Graph Processing) Текстовая аналитика (Text Analysis) Введение в SAC (SAP Analytic Cloud).

### 1.3. Направление, виды воспитательной деятельности и используемые технологии

Направления воспитательной деятельности сопрягаются со всеми результатами обучения компетенций по образовательной программе, их освоение обеспечивается содержанием всех дисциплин модулей.

1.4. Программа дисциплины реализуется на государственном языке Российской Федерации .

## 2. УЧЕБНО-МЕТОДИЧЕСКОЕ И ИНФОРМАЦИОННОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ

### Инструменты Business Intelligence

#### Электронные ресурсы (издания)

1. Бендерская, , О. Б.; Бизнес-аналитика : учебное пособие.; Белгородский государственный технологический университет им. В.Г. Шухова, ЭБС АСВ, Белгород; 2017; <http://www.iprbookshop.ru/92242.html> (Электронное издание)
2. Ушаков, Р. М.; Технология Big Data в деятельности органов исполнительной власти: проблемы и перспективы административно-правового регулирования : студенческая научная работа.; б.и., Саратов; 2021; <https://biblioclub.ru/index.php?page=book&id=617866> (Электронное издание)
3. Воронова, , Л. И.; Big Data. Методы и средства анализа : учебное пособие.; Московский технический университет связи и информатики, Москва; 2016; <http://www.iprbookshop.ru/61463.html> (Электронное издание)
4. Воскобойников, , Ю. Е.; Обработка и анализ экспериментальных данных в пакетах MathCAD и Excel : учебное пособие.; Новосибирский государственный архитектурно-строительный университет (Сибстрин), ЭБС АСВ, Новосибирск; 2020; <http://www.iprbookshop.ru/107639.html> (Электронное издание)
5. Железнов, , М. М.; Методы и технологии обработки больших данных : учебно-методическое пособие.; МИСИ-МГСУ, ЭБС АСВ, Москва; 2020; <http://www.iprbookshop.ru/101802.html> (Электронное издание)

издание)

6. Воронов, В. И.; Data Mining - технологии обработки больших данных : учебное пособие.; Московский технический университет связи и информатики, Москва; 2018; <http://www.iprbookshop.ru/81324.html> (Электронное издание)

7. Чубукова, И. А.; Data Mining : учебное пособие.; Интернет-Университет Информационных Технологий (ИНТУИТ)|Бином. Лаборатория знаний, Москва; 2008; <https://biblioclub.ru/index.php?page=book&id=233055> (Электронное издание)

### **Профессиональные базы данных, информационно-справочные системы**

1. Applied Science & Technology Source. EBSCO publishing <http://search.ebscohost.com>

2. Wiley Online Library <http://onlinelibrary.wiley.com/>

3. Гугл Академия <https://scholar.google.ru/>

### **Материалы для лиц с ОВЗ**

Весь контент ЭБС представлен в виде файлов специального формата для воспроизведения синтезатором речи, а также в тестовом виде, пригодном для прочтения с использованием экранной лупы и настройкой контрастности.

### **Базы данных, информационно-справочные и поисковые системы**

1. Academic Search Ultimate EBSCO publishing – <http://search.ebscohost.com>

2. eBook Collections Springer Nature – <https://link.springer.com/>

3. Электронный научный архив УрФУ <https://elar.urfu.ru/>

4. Зональная научная библиотека (УрФУ) - <http://lib.urfu.ru/>

5. Портал информационно-образовательных ресурсов УрФУ <https://study.urfu.ru/>

6. Электронно-библиотечная система «Лань» – <https://e.lanbook.com/>

7. Университетская библиотека ONLINE – <https://biblioclub.ru/>

8. Электронно-библиотечная система "Библиокомплектатор" (IPRbooks) <http://www.bibliocomplectator.ru/available>

9. Электронные информационные ресурсы Российской государственной библиотеки <https://www.rsl.ru/>

10. Научная электронная библиотека «КиберЛенинка» <https://cyberleninka.ru/>

## **3. МАТЕРИАЛЬНО-ТЕХНИЧЕСКОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ**

### **Инструменты Business Intelligence**

**Сведения об оснащённости дисциплины специализированным и лабораторным оборудованием и программным обеспечением**

Таблица 3.1

№ п/п	Виды занятий	Оснащенность специальных помещений и помещений для самостоятельной работы	Перечень лицензионного программного обеспечения
1	Лекции	<p>Мебель аудиторная с количеством рабочих мест в соответствии с количеством студентов</p> <p>Рабочее место преподавателя</p> <p>Доска аудиторная</p> <p>Оборудование, соответствующее требованиям организации учебного процесса в соответствии с санитарными правилами и нормами</p> <p>Подключение к сети Интернет</p> <p>Мультимедийный проектор с экраном.</p>	Office 365 EDUA3 ShrdSvr ALNG SubsVL MVL PerUsr B Faculty EES
2	Практические занятия	<p>Мебель аудиторная с количеством рабочих мест в соответствии с количеством студентов</p> <p>Рабочее место преподавателя</p> <p>Доска аудиторная</p> <p>Периферийное устройство</p> <p>Персональные компьютеры по количеству обучающихся</p> <p>Оборудование, соответствующее требованиям организации учебного процесса в соответствии с санитарными правилами и нормами</p> <p>Подключение к сети Интернет</p> <p>Мультимедийный проектор с экраном.</p>	<p>Adobe Acrobat Professional 2017 Multiple Platforms</p> <p>Office 365 EDUA3 ShrdSvr ALNG SubsVL MVL PerUsr B Faculty EES</p> <p>В курсе используется бесплатно распространяемое программное обеспечение:</p> <p>Tableau - <a href="https://sso.online.tableau.com/">https://sso.online.tableau.com/</a></p> <p>Платформы для создания опросов: Google Forms - <a href="https://docs.google.com/forms/u/0/">https://docs.google.com/forms/u/0/</a></p> <p>Mentimeter - <a href="https://www.mentimeter.com/">https://www.mentimeter.com/</a></p>

**РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ**  
**Обработка больших объемов данных**

Рабочая программа дисциплины составлена авторами:

<b>№ п/п</b>	<b>Фамилия Имя Отчество</b>	<b>Ученая степень, ученое звание</b>	<b>Должность</b>	<b>Подразделение</b>
1	Корелин Иван Андреевич	кандидат технических наук	Доцент	Базовая кафедра "Автоматизация финансовых систем"

**Рекомендовано учебно-методическим советом института** Радиоэлектроники и информационных технологий - РТФ

Протокол № 4 от 28.05.2021 г.



# 1. СОДЕРЖАНИЕ И ОСОБЕННОСТИ РЕАЛИЗАЦИИ ДИСЦИПЛИНЫ

Авторы:

- Корелин Иван Андреевич, Доцент, Базовая кафедра "Автоматизация финансовых систем"

## 1.1. Технологии реализации, используемые при изучении дисциплины модуля

- Традиционная (репродуктивная) технология
- Разноуровневое (дифференцированное) обучение
  - Базовый уровень

*\*Базовый I уровень – сохраняет логику самой науки и позволяет получить упрощенное, но верное и полное представление о предмете дисциплины, требует знание системы понятий, умение решать проблемные ситуации. Освоение данного уровня результатов обучения должно обеспечить формирование запланированных компетенций и позволит обучающемуся на минимальном уровне самостоятельности и ответственности выполнять задания;*

*Продвинутый II уровень – углубляет и обогащает базовый уровень как по содержанию, так и по глубине проработки материала дисциплины. Это происходит за счет включения дополнительной информации. Данный уровень требует умения решать проблемы в рамках курса и смежных курсов посредством самостоятельной постановки цели и выбора программы действий. Освоение данного уровня результатов обучения позволит обучающемуся повысить уровень самостоятельности и ответственности до творческого применения знаний и умений.*

## 1.2. Содержание дисциплины

Таблица 1.1

Код раздела, темы	Раздел, тема дисциплины*	Содержание
1	Технологии обработки больших данных	Роль и место профессии дата инженера. Основные требуемые навыки. Архитектура центров обработки данных, кластеры для параллельных и распределенных вычислений. Экосистема для распределенного хранения и обработки больших объемов данных: Apache Hadoop, HDFS. Распределённая файловая система Hadoop hdfs и её архитектура. Подход MapReduce на примере Hadoop MapReduce.
2	Основы языка Scala	Понятие функционального программирования. ООП в Scala. Соответствие образцу. Синтаксические конструкции языка. Иммутабельность и единообразие типов коллекций.
3	Фреймворк обработки больших объёмов данных Apache Spark	Распределенная обработка данных в Apache Spark. Архитектура Apache Spark: Resilient Distributed Dataset (RDD), действия трансформации. Сбор данных и формирование набора данных для систем машинного обучения. Методы очистки и подготовки данных. Очистка и подготовка данных. Разметка данных.
4	Кадры данных в Apache Spark	Общедоступные платформы для хранения данных. Варианты для работы с данными в различных форматах: файлы CSV, JSON, HTML. Работа с базами данных. Работа с данными с использованием Spark DataFrame. Источники данных для Spark DataFrame. Обработка данных в Spark DataFrame.

5	Apache Spark SQL	Использование SQL в Spark DataFrame. Форматы хранения больших данных и работа с ними: Parquet, Avro. Графы знаний. Графы вычислений. Производительность в Spark.
6	Apache Spark MLlib	Автоматизация задач машинного обучения с использованием Apache Spark.

### 1.3. Направление, виды воспитательной деятельности и используемые технологии

Направления воспитательной деятельности сопрягаются со всеми результатами обучения компетенций по образовательной программе, их освоение обеспечивается содержанием всех дисциплин модулей.

1.4. Программа дисциплины реализуется на государственном языке Российской Федерации .

## 2. УЧЕБНО-МЕТОДИЧЕСКОЕ И ИНФОРМАЦИОННОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ

### Обработка больших объемов данных

#### Электронные ресурсы (издания)

1. Бутаков, Н. А.; Обработка больших данных с Apache Spark : учебно-методическое пособие.; Университет ИТМО, Санкт-Петербург; 2019; <https://biblioclub.ru/index.php?page=book&id=566771> (Электронное издание)
2. Воронов, В. И.; Data Mining - технологии обработки больших данных : учебное пособие.; Московский технический университет связи и информатики, Москва; 2018; <http://www.iprbookshop.ru/81324.html> (Электронное издание)
3. Железнов, М. М.; Методы и технологии обработки больших данных : учебно-методическое пособие.; МИСИ-МГСУ, ЭБС АСВ, Москва; 2020; <http://www.iprbookshop.ru/101802.html> (Электронное издание)
4. Кей, Киселев, А. Н.; Scala для нетерпеливых; Профобразование, Саратов; 2019; <http://www.iprbookshop.ru/87976.html> (Электронное издание)

### Профессиональные базы данных, информационно-справочные системы

1. Applied Science & Technology Source. EBSCO publishing <http://search.ebscohost.com>
2. Wiley Online Library <http://onlinelibrary.wiley.com/>
3. Гугл Академия <https://scholar.google.ru/>

### Материалы для лиц с ОВЗ

Весь контент ЭБС представлен в виде файлов специального формата для воспроизведения синтезатором речи, а также в тестовом виде, пригодном для прочтения с использованием экранной лупы и настройкой контрастности.

### Базы данных, информационно-справочные и поисковые системы

1. Academic Search Ultimate EBSCO publishing – <http://search.ebscohost.com>
2. eBook Collections Springer Nature – <https://link.springer.com/>

3. Электронный научный архив УрФУ <https://elar.urfu.ru/>
4. Зональная научная библиотека (УрФУ) - <http://lib.urfu.ru/>
5. Портал информационно-образовательных ресурсов УрФУ <https://study.urfu.ru/>
6. Электронно-библиотечная система «Лань» – <https://e.lanbook.com/>
7. Университетская библиотека ONLINE – <https://biblioclub.ru/>
8. Электронно-библиотечная система "Библиокомплектатор" (IPRbooks) <http://www.bibliocomplectator.ru/available>
9. Электронные информационные ресурсы Российской государственной библиотеки <https://www.rsl.ru/>
10. Научная электронная библиотека «КиберЛенинка» <https://cyberleninka.ru/>

### 3. МАТЕРИАЛЬНО-ТЕХНИЧЕСКОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ

#### Обработка больших объемов данных

#### Сведения об оснащённости дисциплины специализированным и лабораторным оборудованием и программным обеспечением

Таблица 3.1

№ п/п	Виды занятий	Оснащённость специальных помещений и помещений для самостоятельной работы	Перечень лицензионного программного обеспечения
1	Лекции	<p>Мебель аудиторная с количеством рабочих мест в соответствии с количеством студентов</p> <p>Рабочее место преподавателя</p> <p>Доска аудиторная</p> <p>Оборудование, соответствующее требованиям организации учебного процесса в соответствии с санитарными правилами и нормами</p> <p>Подключение к сети Интернет</p> <p>Мультимедийный проектор с экраном</p>	<p>Adobe Acrobat Professional 2017 Multiple Platforms</p> <p>Office 365 EDUA3 ShrdSvr ALNG SubsVL MVL PerUsr B Faculty EES</p> <p>В курсе используется бесплатно распространяемое программное обеспечение:</p> <p>Python – <a href="https://www.python.org/">https://www.python.org/</a></p> <p>Веб - среда разработки для языка программирования Python: google colab - <a href="https://colab.research.google.com/">https://colab.research.google.com/</a></p> <p>Anaconda solution - <a href="https://www.anaconda.com/">https://www.anaconda.com/</a></p> <p>Система распределенной обработки данных Apache Spark – <a href="https://spark.apache.org/">https://spark.apache.org/</a></p> <p>Кластер HDInsight от Microsoft Azure</p>

2	Практические занятия	<p>Мебель аудиторная с количеством рабочих мест в соответствии с количеством студентов</p> <p>Рабочее место преподавателя</p> <p>Доска аудиторная</p> <p>Периферийное устройство</p> <p>Персональные компьютеры по количеству обучающихся</p> <p>Оборудование, соответствующее требованиям организации учебного процесса в соответствии с санитарными правилами и нормами</p> <p>Подключение к сети Интернет</p> <p>Мультимедийный проектор с экраном</p>	<p>Office 365 EDUA3 ShrdSvr ALNG SubsVL MVL PerUsr B Faculty EES</p> <p>В курсе используется бесплатно распространяемое программное обеспечение:</p> <p>Python – <a href="https://www.python.org/">https://www.python.org/</a></p> <p>Веб - среда разработки для языка программирования Python: google colab - <a href="https://colab.research.google.com/">https://colab.research.google.com/</a></p> <p>Anaconda solution - <a href="https://www.anaconda.com/">https://www.anaconda.com/</a></p> <p>Система распределенной обработки данных Apache Spark – <a href="https://spark.apache.org/">https://spark.apache.org/</a></p> <p>Кластер HDInsight от Microsoft Azure</p>