

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
2021 г.

Мас



РАБОЧАЯ ПРОГРАММА МОДУЛЯ

Код модуля
М.1.5

Модуль
Базы данных

Екатеринбург, 2021

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

Области образования, в рамках которых реализуется модуль образовательной программы по СУОС УрФУ:

№ п/п	Перечень областей образования, для которых разработан СУОС УрФУ	Уровень подготовки
1.	Инженерное дело, технологии и технические науки	магистратура

Программа модуля составлена авторами:

№ п/п	Фамилия Имя Отчество	Ученая степень, ученое звание	Должность	Подразделение
1	Созыкин Андрей Владимирович	Кандидат технических наук, нет	Доцент	Кафедра информационных технологий и систем управления, ИРИТ-РТФ, УрФУ
2	Медведев Александр Николаевич	Кандидат технических наук, старший научный сотрудник	Доцент	Базовая кафедра "Аналитика больших данных и методы видеонализа", ИРИТ-РТФ, УрФУ
3	Корелин Иван Андреевич	Кандидат технических наук, нет	Старший преподаватель	Учебно-научный центр "Информационная безопасность", ИРИТ-РТФ, УрФУ
4	Бурак Илья Александрович	-	Главный инженер по разработке	ПАО Сбербанк

Рекомендовано учебно-методическим советом института радиоэлектроники и информационных технологий - РТФ

Протокол № 4 от 28.05.2021 г.

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», «Обработка больших объемов данных» посвященных полному циклу работы с данными: работа с реляционными СУБД, визуализация и бизнес-анализ данных, подготовка данных для моделей машинного обучения и систем бизнес-отчетности и принятия решений. Рассматриваются особенности работы с данными в различных форматах на языках SQL, Python, scala, Java, отдельное внимание уделяется языку PL\SQL от Oracle. Уделяется внимание инструментам и технологиям загрузки данных из различных источников. Подробно изучаются методы очистки данных и соответствующие библиотеки. Рассматриваются технологии Oracle, Hadoop, Spark.

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 и часах
1.	Базы данных	6 з.е. / 216 ч.
2.	Инструменты Business Intelligence	6 з.е. / 216 ч.
3.	Обработка больших объемов данных	6 з.е. / 216 ч.
ИТОГО по модулю:		18 з.е. / 648 ч.

1.3. Последовательность освоения модуля в образовательной программе

Пререквизиты модуля	<i>отсутствуют</i>
Постреквизиты и корреквизиты модуля	<i>отсутствуют</i>

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

Таблица 2.1

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

		<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>
	<p>ПК-7 - Способен разрабатывать системы управления базами данных</p>	<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	<p>ПК-4 - Способен осуществлять управление развитием баз данных</p>	<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>
	<p>ПК-7 - Способен разрабатывать системы управления базами данных</p>	<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</p>

		<p>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 дисциплине модуля может осуществляться в очной форме.

2. СОДЕРЖАНИЕ И ОБЕСПЕЧЕНИЕ РЕАЛИЗАЦИИ ДИСЦИПЛИН МОДУЛЯ

**ПРОГРАММА МОДУЛЯ
БАЗЫ ДАННЫХ**

**РАЗДЕЛ 2. СОДЕРЖАНИЕ И ОБЕСПЕЧЕНИЕ РЕАЛИЗАЦИИ ДИСЦИПЛИН
МОДУЛЯ**

**РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ 1
БАЗЫ ДАННЫХ**

Рабочая программа дисциплины составлена авторами:

№ п/п	Фамилия Имя Отчество	Ученая степень, ученое звание	Должность	Подразделение
1	Созыкин Андрей Владимирович	Кандидат технических наук, нет	Доцент	Кафедра информационных технологий и систем управления, ИРИТ-РТФ, УрФУ
2	Бурак Илья Александрович	-	Главный инженер по разработке	ПАО Сбербанк

Рекомендовано учебно-методическим советом института радиоэлектроники и информационных технологий - РТФ

Протокол № 4 от 28.05.2021 г.

2. СОДЕРЖАНИЕ И ОСОБЕННОСТИ РЕАЛИЗАЦИИ ДИСЦИПЛИНЫ Базы данных

2.1. Технологии обучения, используемые при изучении дисциплины модуля

– Традиционная (репродуктивная) технология;

2.2. Содержание дисциплины

Таблица 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

2.3. Программа дисциплины реализуется на государственном языке Российской Федерации.

2.4. УЧЕБНО-МЕТОДИЧЕСКОЕ И ИНФОРМАЦИОННОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ Базы данных

Электронные ресурсы (издания)

1. Аврунев, О. Е. Модели баз данных : учебное пособие : О. Е. Аврунев, В. М. Стасышин. – Новосибирск : Новосибирский государственный технический университет, 2018. – 124 с. : ил., табл. – Режим доступа: по подписке. – URL: <https://biblioclub.ru/index.php?page=book&id=575324> – Библиогр. в кн. – ISBN 978-5-7782-3749-0. – Текст : электронный.
2. Зудилова, Т. В. SQL и PL/SQL для разработчиков СУБД Oracle / Т. В. Зудилова, С. Е. Иванов, С. Э. Хоружников. — Санкт-Петербург : Университет ИТМО, 2012. — 73 с. — Текст : электронный // Электронно-библиотечная система IPR BOOKS : [сайт]. — URL: <https://www.iprbookshop.ru/65745.html> — Режим доступа: для авторизир. пользователей
3. Кузнецов, С. Д. Введение в модель данных SQL : учебное пособие / С. Д. Кузнецов. — 3-е изд. — Москва : Интернет-Университет Информационных Технологий (ИНТУИТ), Ай Пи Ар Медиа, 2021. — 350 с. — ISBN 978-5-4497-0873-1. — Текст : электронный // Электронно-библиотечная система IPR BOOKS : [сайт]. — URL: <https://www.iprbookshop.ru/101995.html>
4. Пржиялковский, В. В. Введение в Oracle SQL : учебное пособие / В. В. Пржиялковский. — 3-е изд. — Москва : Интернет-Университет Информационных Технологий (ИНТУИТ), Ай Пи Ар Медиа, 2020. — 336 с. — ISBN 978-5-4497-0543-3.

— Текст : электронный // Электронно-библиотечная система IPR BOOKS : [сайт]. — URL: <https://www.iprbookshop.ru/94846.html>

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/>

2.5. МАТЕРИАЛЬНО-ТЕХНИЧЕСКОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ Базы данных

Сведения об оснащённости дисциплины специализированным и лабораторным оборудованием и программным обеспечением

Таблица 2

№ п/п	Виды занятий	Оснащённость специальных помещений и помещений для самостоятельной работы	Перечень лицензионного программного обеспечения. Реквизиты подтверждающего документа
1	Лекции; Практические занятия.	Компьютерный класс. Мультимедийный проектор с экраном; Сетевое оборудование.; Локальная сеть с выходом в глобальную сеть Internet.	В курсе используется бесплатно распространяемое программное обеспечение: Oracle Database XE, Oracle SQL Developer.

**ПРОГРАММА МОДУЛЯ
БАЗЫ ДАННЫХ**

**РАЗДЕЛ 2. СОДЕРЖАНИЕ И ОБЕСПЕЧЕНИЕ РЕАЛИЗАЦИИ ДИСЦИПЛИН
МОДУЛЯ**

**РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ 2
Инструменты Business Intelligence**

Рабочая программа дисциплины составлена авторами:

№ п/п	Фамилия Имя Отчество	Ученая степень, ученое звание	Должность	Подразделение
1	Медведев Александр Николаевич	Кандидат технических наук, старший научный сотрудник	Доцент	Базовая кафедра «Аналитика больших данных и методы видеоанализа», ИРИТ-РТФ, УрФУ

Рекомендовано учебно-методическим советом института радиоэлектроники и информационных технологий - РТФ

Протокол № 4 от 28.05.2021 г.

2. СОДЕРЖАНИЕ И ОСОБЕННОСТИ РЕАЛИЗАЦИИ ДИСЦИПЛИНЫ Инструменты Business Intelligence

2.1. Технологии обучения, используемые при изучении дисциплины модуля

– Традиционная (репродуктивная) технология;

2.2. Содержание дисциплины

Таблица 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 наиболее важной информации.	Анализ полученной неструктурированной информации методами Power BI, Excel и Google Sheets
P7	Метод презентации аналитических дашбордов широкой аудитории	Разбор методик презентаций проектов AIDA и Elevator Speech Практика в представлении проекта с помощью методики AIDA
P8	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)

2.3. Программа дисциплины реализуется на государственном языке Российской Федерации.

2.4. УЧЕБНО-МЕТОДИЧЕСКОЕ И ИНФОРМАЦИОННОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ Инструменты Business Intelligence

Электронные ресурсы (издания)

1. Бернанд, Марр Ключевые инструменты бизнес-аналитики. 67 инструментов, которые должен знать каждый менеджер / Марр Бернанд ; перевод В. Н. Егоров. — Москва : Лаборатория знаний, 2018. — 337 с. — ISBN 978-5-00101-610-6. — Текст : электронный // Электронно-библиотечная система IPR BOOKS : [сайт]. — URL: <https://www.iprbookshop.ru/89019.html>
2. Бендерская, О. Б. Бизнес-аналитика : учебное пособие / О. Б. Бендерская. — Белгород : Белгородский государственный технологический университет им. В.Г. Шухова, ЭБС АСВ, 2017. — 162 с. — Текст : электронный // Электронно-библиотечная система IPR BOOKS : [сайт]. — URL: <https://www.iprbookshop.ru/92242.html>
3. Билл, Фрэнк Революция в аналитике: Как в эпоху Big Data улучшить ваш бизнес с помощью операционной аналитики / Фрэнк Билл ; перевод И. Евстигнеева ; под редакцией В. Мылова. — Москва : Альпина Паблишер, 2020. — 320 с. — ISBN 978-5-9614-5302-7. — Текст : электронный // Электронно-библиотечная система IPR BOOKS : [сайт]. — URL: <https://www.iprbookshop.ru/93032.html>
4. Воскобойников, Ю. Е. Обработка и анализ экспериментальных данных в пакетах MathCAD и Excel : учебное пособие / Ю. Е. Воскобойников. — Новосибирск : Новосибирский государственный архитектурно-строительный университет (Сибстрин), ЭБС АСВ, 2020. — 161 с. — ISBN 978-5-7795-0906-0. — Текст : электронный // Электронно-библиотечная система IPR BOOKS : [сайт]. — URL: <https://www.iprbookshop.ru/107639.html>
5. Железнов, М. М. Методы и технологии обработки больших данных : учебно-методическое пособие / М. М. Железнов. — Москва : МИСИ-МГСУ, ЭБС АСВ, 2020. — 46 с. — ISBN 978-5-7264-2193-3. — Текст : электронный // Электронно-библиотечная система IPR BOOKS : [сайт]. — URL: <https://www.iprbookshop.ru/101802.html>
6. Воронов, В. И. Data Mining - технологии обработки больших данных : учебное пособие / В. И. Воронов, Л. И. Воронова, В. А. Усачев. — Москва : Московский технический университет связи и информатики, 2018. — 47 с. — Текст : электронный // Электронно-библиотечная система IPR BOOKS : [сайт]. — URL: <https://www.iprbookshop.ru/81324.html>
7. Чубукова, И. А. Data Mining : учебное пособие / И. А. Чубукова. — 3-е изд. — Москва, Саратов : Интернет-Университет Информационных Технологий (ИНТУИТ), Ай Пи Ар Медиа, 2020. — 469 с. — ISBN 978-5-4497-0289-0. — Текст : электронный // Электронно-библиотечная система IPR BOOKS : [сайт]. — URL: <https://www.iprbookshop.ru/89404.html>
8. David Slager. Essential Excel 2016 Apress, Berkeley, CA, XXI, 673 p., 2016, DOI <https://ezproxy.urfu.ru:3055/10.1007/978-1-4842-2161-7>, ISBN 978-1-4842-2160-0.
9. Abbott Katz. Beginning Microsoft Excel 2010, Apress, 408 p., 2010, DOI <https://ezproxy.urfu.ru:3055/10.1007/978-1-4302-2956-8>, ISBN 978-1-4302-2955-1.
10. Neil Dunlop. Beginning Big Data with Power BI and Excel 2013, Apress, Berkeley, CA, XIX, 246 p., 2015, DOI <https://ezproxy.urfu.ru:3055/10.1007/978-1-4842-0529-7>, ISBN 978-1-4842-0530-3.
11. Thomas J Quirk. Excel 2010 for Business Statistics, Springer, New York, NY, XV, 252 p., 2011, DOI <https://ezproxy.urfu.ru:3055/10.1007/978-1-4419-9934-4>, ISBN 978-1-4419-9933-7.

12. Mercurio R., Merrill B. Power BI. In: Beginning Microsoft 365 Collaboration Apps. Apress, Berkeley, CA, (2021), https://ezproxy.urfu.ru:3055/10.1007/978-1-4842-6936-7_12.
13. Adam Aspin. Pro Power BI Desktop, Apress, Berkeley, CA, XXXVI, 761 p., 2018, DOI <https://ezproxy.urfu.ru:3055/10.1007/978-1-4842-3210-1>, ISBN 978-1-4842-3209-5.
14. Adam Aspin. High Impact Data Visualization with Power View, Power Map, and Power BI, Apress, Berkeley, CA, XXVI, 552 p., 2014, DOI <https://ezproxy.urfu.ru:3055/10.1007/978-1-4302-6617-4>, ISBN 978-1-4302-6616-7.
15. Dan Clark. Beginning Power BI with Excel 2013, Apress, Berkeley, CA, XIV, 324 p., 2013, DOI <https://ezproxy.urfu.ru:3055/10.1007/978-1-4302-6446-0>, ISBN 978-1-4302-6445-3.
16. Ryan Wade. Advanced Analytics in Power BI with R and Python, Apress, Berkeley, CA, XLVI, 391, 2020, DOI <https://ezproxy.urfu.ru:3055/10.1007/978-1-4842-5829-3>, ISBN 978-1-4842-5828-6.
17. Adam Aspin. High Impact Data Visualization in Excel with Power View, 3D Maps, Get & Transform and Power BI, Apress, Berkeley, CA, XXXII, 599, 2016, DOI <https://ezproxy.urfu.ru:3055/10.1007/978-1-4842-2400-7>, ISBN 978-1-4842-2399-4.
18. Kolokolov A.S. Power BI for business users. Basic tips to develop analytical dashboard in Power BI. UdeMy Massive open online course. url: <https://www.udemy.com/course/power-bi-for-business-users/?instructorPreviewMode=guest>
19. Kolokolov A.S. In Pursuit of a New Golden Ratio. Medium, 2020. url: <https://medium.com/nightingale/in-pursuit-of-a-new-golden-ratio-1ad528534222>
20. Kolokolov A.S How happy are data analysts? Visualization of analysts' salary poll. Medium, url: <https://towardsdatascience.com/how-happy-are-data-analysts-visualization-of-analyst-salary-poll-422c1c50a027>
21. Cairo A. Stork A. and Basole R. If Anything on This Graphic Causes Confusion, Discard the Entire Product. IEEE Computer Graphics and Applications. 40. 91-97. 10.1109/MCG.2019.2961716, 2020.
22. Claus O. Wilke. Fundamentals of Data Visualization: A Primer on Making Informative and Compelling Figures, O'Reilly Media, Inc.", 2019. url=<https://books.google.ru/books?id=WmmNDwAAQBAJ>

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/>

2.5. МАТЕРИАЛЬНО-ТЕХНИЧЕСКОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ

Инструменты Business Intelligence

Сведения об оснащённости дисциплины специализированным и лабораторным оборудованием и программным обеспечением

Таблица 2

№ п/п	Виды занятий	Оснащённость специальных помещений и помещений для самостоятельной работы	Перечень лицензионного программного обеспечения. Реквизиты подтверждающего документа
1	Лекции; Практические занятия.	Компьютерный класс. Мультимедийный проектор с экраном.; Сетевое оборудование.; Локальная сеть с выходом в глобальную сеть Internet.	В курсе используется бесплатно распространяемое программное обеспечение: Microsoft Excel Платформы бизнес-аналитики: Microsoft Power BI - https://powerbi.microsoft.com/ru-ru/desktop/ Tableau - https://sso.online.tableau.com/ Платформы для создания опросов: Google Forms - https://docs.google.com/forms/u/0/ Mentimeter - https://www.mentimeter.com/

**ПРОГРАММА МОДУЛЯ
БАЗЫ ДАННЫХ**

**РАЗДЕЛ 2. СОДЕРЖАНИЕ И ОБЕСПЕЧЕНИЕ РЕАЛИЗАЦИИ ДИСЦИПЛИН
МОДУЛЯ**

**РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ 3
ОБРАБОТКА БОЛЬШИХ ОБЪЕМОВ ДАННЫХ**

Рабочая программа дисциплины составлена авторами:

№ п/п	Фамилия Имя Отчество	Ученая степень, ученое звание	Должность	Подразделение
1	Корелин Иван Андреевич	кандидат технических наук, нет	Старший преподаватель	Учебно-научный центр ”Информационная безопасность”, ИРИТ-РТФ, УрФУ

Рекомендовано учебно-методическим советом института радиоэлектроники и информационных технологий - РТФ

Протокол № 4 от 28.05.2021 г.

2. СОДЕРЖАНИЕ И ОСОБЕННОСТИ РЕАЛИЗАЦИИ ДИСЦИПЛИНЫ Обработка больших объемов данных

2.1. Технологии обучения, используемые при изучении дисциплины модуля

– Традиционная (репродуктивная) технология;

2.2. Содержание дисциплины

Таблица 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.

2.3. Программа дисциплины реализуется на государственном языке Российской Федерации.

2.4. 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 / Н. А. Бутаков, М. В. Петров, Д. Насонов .— Санкт-Петербург : Университет ИТМО, 2019 .— 52 с. : ил. — Библиогр. в кн .— Режим доступа: электронная библиотечная система «Университетская библиотека ONLINE», требуется авторизация — URL:<https://biblioclub.ru/index.php?page=book&id=566771>
2. Воронов, В. И. Data Mining - технологии обработки больших данных : учебное пособие / В. И. Воронов, Л. И. Воронова, В. А. Усачев. — Москва : Московский технический университет связи и информатики, 2018. — 47 с. — Текст : электронный // Электронно-библиотечная система IPR BOOKS : [сайт]. — URL: <https://www.iprbookshop.ru/81324.html>
 3. Железнов, М. М. Методы и технологии обработки больших данных : учебно-методическое пособие / М. М. Железнов. — Москва : МИСИ-МГСУ, ЭБС АСВ, 2020. — 46 с. — ISBN 978-5-7264-2193-3. — Текст : электронный // Электронно-библиотечная система IPR BOOKS : [сайт]. — URL: <https://www.iprbookshop.ru/101802.html>
 4. Кей, Хорстманн Scala для нетерпеливых / Хорстманн Кей ; перевод А. Н. Киселев. — Саратов : Профобразование, 2019. — 414 с. — ISBN 978-5-4488-0434-2. — Текст : электронный // Электронно-библиотечная система IPR BOOKS : [сайт]. — URL: <https://www.iprbookshop.ru/87976.html>
 5. Белов В. А., Никульчев Е. В. ЭКСПЕРИМЕНТАЛЬНАЯ ОЦЕНКА ВРЕМЕННОЙ ЭФФЕКТИВНОСТИ ОБРАБОТКИ БОЛЬШИХ ДАННЫХ В ЗАДАННЫХ ФОРМАТАХ ХРАНЕНИЯ // International Journal of Open Information Technologies. 2021. №9. URL: <https://cyberleninka.ru/article/n/eksperimentalnaya-otsenka-vremennoy-effektivnosti-obrabotki-bolshih-dannyh-v-zadannyh-formatah-hraneniya>.
 6. Манев Дмитрий Валерьевич, Сальников Вячеслав Юрьевич Информационная система обработки и хранения больших объемов измерительных данных // SAEC. 2019. №1. URL: <https://cyberleninka.ru/article/n/informatsionnaya-sistema-obrabotki-i-hraneniya-bolshih-obemov-izmeritelnyh-dannyh>

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/>

2.5. МАТЕРИАЛЬНО-ТЕХНИЧЕСКОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ Обработка больших объемов данных

Сведения об оснащённости дисциплины специализированным и лабораторным оборудованием и программным обеспечением

Таблица 2

№ п/п	Виды занятий	Оснащённость специальных помещений и помещений для самостоятельной работы	Перечень лицензионного программного обеспечения. Реквизиты подтверждающего документа
1	Лекции; Практические занятия.	Компьютерный класс. Мультимедийный проектор с экраном.; Сетевое оборудование.; Локальная сеть с выходом в глобальную сеть Internet.	В курсе используется бесплатно распространяемое программное обеспечение: 1. Python – https://www.python.org/ 2. Веб - среда разработки для языка программирования Python: google colab - https://colab.research.google.com/ 3. Anaconda solution - https://www.anaconda.com/ 4. Система распределенной обработки данных Apache Spark – https://spark.apache.org/ 5. Кластер HDInsight от Microsoft Azure