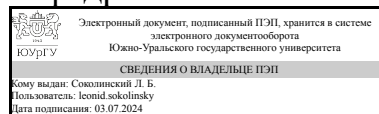


ЮЖНО-УРАЛЬСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ

УТВЕРЖДАЮ:
Заведующий выпускающей
кафедрой



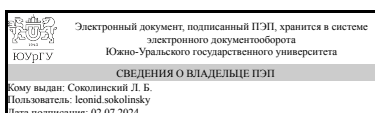
Л. Б. Соколинский

РАБОЧАЯ ПРОГРАММА

дисциплины 1.Ф.П0.08 Основы распределенной обработки данных
для направления 09.03.04 Программная инженерия
уровень Бакалавриат
профиль подготовки Инженерия информационных и интеллектуальных систем
форма обучения очная
кафедра-разработчик Системное программирование

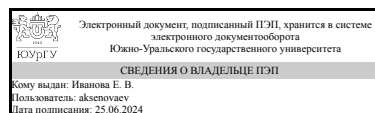
Рабочая программа составлена в соответствии с ФГОС ВО по направлению подготовки 09.03.04 Программная инженерия, утверждённым приказом Минобрнауки от 19.09.2017 № 920

Зав.кафедрой разработчика,
д.физ.-мат.н., проф.



Л. Б. Соколинский

Разработчик программы,
к.физ.-мат.н., доцент



Е. В. Иванова

1. Цели и задачи дисциплины

Целью курса является изучение студентами задач, связанных с распределенным хранением и обработкой больших данных. При изучении этого курса должны быть решены следующие задачи: изучение понятия и проблематики больших данных, способы распределенного хранения и обработки больших данных, хранение и обработка больших данных в экосистеме Hadoop.

Краткое содержание дисциплины

Понятие больших данных. Распределенная обработка больших данных. Основы Hadoop, HDFS, MapReduce. Экосистема Hadoop: Pig, Apache Hive, HBase, Apache Spark, MLlib, Hadoop YARN, Zookeeper, Apache Kafka.

2. Компетенции обучающегося, формируемые в результате освоения дисциплины

Планируемые результаты освоения ОП ВО (компетенции)	Планируемые результаты обучения по дисциплине
ПК-2 Способен разрабатывать компоненты системных программных продуктов на основе соответствующей технической документации	Знает: основы работы компонентов экосистемы Hadoop Умеет: строить программную систему на основе компонентов экосистемы Hadoop для решения поставленной задачи Имеет практический опыт: создания программной системы на основе компонентов экосистемы Hadoop
ПК-7 (ПК-8 модели) Способен разрабатывать системы анализа больших данных	Знает: ПК-8.1. 3-2. Знает принципы работы экосистемы Hadoop, фреймворка SPARK; ПК-8.2. 3-1. Знает принципы и методы анализа больших данных, включая спецификации и стандартизацию метаданных; ПК-8.2. 3-2. Знает устройство и принципы работы систем обработки и анализа больших массивов данных (SQL, NoSQL, Hadoop, ETL); ПК-8.2. 3-3. Знает архитектуру и принципы работы промышленных решений, созданных на основе искусственного интеллекта; Умеет: ПК-8.1. У-4. Умеет использовать шины данных (Apache Kafka); ПК-8.2. У-3. Умеет использовать системы обработки и анализа больших массивов данных (SQL, NoSQL, Hadoop, ETL процессы и инструменты); ПК-8.2. У-4. Умеет использовать технологии науки о данных и больших данных в разработке для решения практических задач промышленности; Имеет практический опыт: обработки и анализа данных в экосистеме Hadoop

3. Место дисциплины в структуре ОП ВО

Перечень предшествующих дисциплин, видов работ учебного плана	Перечень последующих дисциплин, видов работ
---	---

Базы данных, Технологии аналитической обработки информации, Архитектура ЭВМ, Современные языки программирования систем искусственного интеллекта, Администрирование ОС Linux, Основы интеллектуального анализа данных	Не предусмотрены
--	------------------

Требования к «входным» знаниям, умениям, навыкам студента, необходимым при освоении данной дисциплины и приобретенным в результате освоения предшествующих дисциплин:

Дисциплина	Требования
Архитектура ЭВМ	Знает: основные положения и концепции в области архитектуры ЭВМ, базовые принципы проектирования системного ПО, типы архитектур ЭВМ, требования к системному и прикладному ПО, понятие архитектуры ЭВМ, способы представления данных в ЭВМ, принципы организации вычислений Умеет: решать стандартные задачи в профессиональной деятельности с учетом способов представления и обработки данных в ЭВМ, проектировать ПО с учетом принципов организации ЭВМ, разрабатывать алгоритмические и программные решения с использованием низкоуровневых языков программирования Имеет практический опыт: разработки программ на низкоуровневых языках программирования с учетом способов представления и обработки данных в ЭВМ, проектирования системного ПО с учетом принципов организации ЭВМ, системного программирования с использованием низкоуровневых языков программирования
Современные языки программирования систем искусственного интеллекта	Знает: инструментальные средства для решения задач машинного обучения , ПК-8.1. 3-4. Знает предметно-ориентированные языки;, ПК-2.1. 3-1. Знает основные программные платформы и компоненты систем искусственного интеллекта: механизмы логического вывода (рассуждений), объяснений, приобретения знаний, интеллектуальных интерфейсов, принципы Data Ops и Dev Ops;ПК-2.2. 3-1. Знает современные языки программирования, библиотеки и программные платформы для функционального, логического, объектно-ориентированного программирования, приложений систем искусственного интеллекта (Python, R, C++, C#); Умеет: ПК-5.1. У-1. Умеет проводить сравнительный анализ и осуществлять выбор инструментальных средств для решения задач машинного обучения;, ПК-2.1. У-1. Умеет настраивать основные программные платформы и компоненты систем искусственного

	<p>интеллекта: механизмов логического вывода (рассуждений), объяснений, приобретения знаний, интеллектуальных интерфейсов на особенности проблемной области, участвует в их разработке; ПК-2.2. У-1. Умеет разрабатывать программные приложения систем искусственного интеллекта с использованием программных платформ функционального, логического, объектно-ориентированного программирования (Python, R, C++, C#); Имеет практический опыт: разработки приложений для машинного обучения на языках программирования систем искусственного интеллекта (Python, R, C++, C#), создания приложений искусственного интеллекта с использованием современных языков программирования, библиотек и программных платформ</p>
Базы данных	<p>Знает: ПК-8.1. З-3. Знает устройство интерфейсов между реляционными SQL-хранилищами данных и нереляционными NoSQL-хранилищами данных; основы работы современных систем управления базами данных, основы устройства систем баз данных Умеет: ПК-8.1. У-5. Умеет использовать языки запросов, в том числе нереляционных, для поддержки различных типов данных (например, XML, RTF, JSON, мультимедиа) и операций с большими данными (например, матричные операции); создавать реляционные и нереляционные базы данных и запросы к ним, устанавливать и настраивать реляционные и нереляционные системы баз данных Имеет практический опыт: написания запросов к реляционным и нереляционным большим базам данных, разработки реляционных и нереляционных баз данных, инсталляции систем баз данных</p>
Администрирование ОС Linux	<p>Знает: основные принципы устройства файловой системы в Linux, межпроцессное и многопоточное взаимодействие, ПК-2.1.3-1. Знает основные программные платформы и компоненты систем искусственного интеллекта: механизмы логического вывода (рассуждений), объяснений, приобретения знаний, интеллектуальных интерфейсов, принципы Data Ops и Dev Ops, основные принципы устройства и администрирования ОС семейства Linux, принципы разработки системных утилит в Linux Умеет: разрабатывать системные решения обработки файлов в Linux, реализацию многопоточных приложений, клиент-серверных приложений в Linux, ПК-2.1. У-1. Умеет настраивать основные программные платформы и компоненты систем искусственного интеллекта: механизмов логического вывода (рассуждений), объяснений, приобретения</p>

	<p>знаний, интеллектуальных интерфейсов на особенности проблемной области, участвует в их разработке;,, выполнять задачи администрирования ОС семейства Linux, реализовывать системные скрипты для решения задач профессиональной деятельности Имеет практический опыт: работы с основными утилитами командной строки в Linux</p>
<p>Основы интеллектуального анализа данных</p>	<p>Знает: ПК-9.3. 3-1. Знает фундаментальные правила построения рекомендательных систем и систем поддержки принятия решений, основанных на интеллектуальных принципах, методы и подходы к планированию и реализации проектов по созданию систем искусственного интеллекта на основе сквозной цифровой субтехнологии «Рекомендательные системы и системы поддержки принятия решений»; ПК-8.1. 3-1. Знает общедоступные репозитории и специализированные библиотеки, содержащие наборы больших данных;ПК-8.2. 3-4. Знает методы и технологии машинного обучения на больших данных;,, ПК-4.1. 3-1. Знает принципы и методы машинного обучения, типы и классы задач машинного обучения, методологию ML Ops;ПК-4.2. 3-1. Знает методы и критерии оценки качества моделей машинного обучения;,, ПК-5.3. 3-1. Знает принципы построения систем искусственного интеллекта , методы и подходы к планированию и реализации проектов по созданию систем искусственного интеллекта с применением машинного обучения;постановку и методы решения основных задач интеллектуального анализа данных (поиск шаблонов, классификация, кластеризация) Умеет: ПК-9.3. У-1. Умеет применять методы и подходы к планированию и реализации проектов по созданию и поддержке системы искусственного интеллекта на основе сквозной цифровой субтехнологии «Рекомендательные системы и системы поддержки принятия решений»;,, ПК-8.1. У-1. Умеет настраивать и оптимизировать конфигурацию программного и аппаратного обеспечения с целью интеграции больших данных;ПК-8.1. У-2. Умеет разрабатывать программное обеспечение для очистки и валидации наборов больших данных;ПК-8.2. У-1. Умеет разрабатывать программное обеспечение для анализа больших данных;ПК-8.2. У-2. Умеет разрабатывать программные и технические средства визуализации больших данных и результатов их анализа;ПК-8.2. У-5. Умеет описывать и управлять качеством и достоверностью больших данных;,, ПК-4.1. У-1. Умеет сопоставить задачам предметной области классы задач машинного обучения;ПК-4.1. У-2. Умеет использовать</p>

	<p>статистические методы анализа данных при решении задач машинного обучения; ПК-4.2. У-1. Умеет определять критерии и метрики оценки результатов моделирования при построении систем искусственного интеллекта в исследуемой области; ПК-5.2. У-2. Умеет планировать и выполнять машинные эксперименты, оценивать точность и качество построенных моделей; ПК-5.3. У-1. Умеет решать задачи по выполнению коллективной проектной деятельности для создания, поддержки и использования системы искусственного интеллекта с применением машинного обучения и массово параллельных вычислений для ускорения машинного обучения; Имеет практический опыт: разработки программных компонент для извлечения и подготовки больших данных для интеллектуального анализа, анализа требований и определения необходимых классов задач для реализации приложений машинного обучения; определения метрик и критериев качества оценки моделей машинного обучения, разработки моделей машинного обучения для решения основных задач интеллектуального анализа данных (поиск шаблонов, классификация, кластеризация) и проведения вычислительных экспериментов по оценке точности и качества построенных моделей</p>
<p>Технологии аналитической обработки информации</p>	<p>Знает: общедоступные репозитории и специализированные библиотеки, содержащие наборы больших данных, методы и критерии оценки качества моделей машинного обучения, постановку базовых задач интеллектуального анализа данных (поиск шаблонов, классификация, кластеризация) и базовые методы их решения Умеет: настраивать и оптимизировать конфигурацию программного и аппаратного обеспечения с целью интеграции больших данных, сопоставить задачам предметной области классы задач машинного обучения, планировать и выполнять машинные эксперименты, оценивать точность и качество построенных моделей Имеет практический опыт: разработки программных компонент для извлечения и подготовки больших данных для аналитической обработки информации, анализа требований и идентификации классов задач для реализации приложений машинного обучения, разработки приложений для аналитической обработки информации с помощью современных инструментальных средств</p>

4. Объём и виды учебной работы

Общая трудоемкость дисциплины составляет 4 з.е., 144 ч., 58,5 ч. контактной работы

Вид учебной работы	Всего часов	Распределение по семестрам в часах	
		Номер семестра	
		8	
Общая трудоёмкость дисциплины	144	144	
<i>Аудиторные занятия:</i>	48	48	
Лекции (Л)	24	24	
Практические занятия, семинары и (или) другие виды аудиторных занятий (ПЗ)	24	24	
Лабораторные работы (ЛР)	0	0	
<i>Самостоятельная работа (СРС)</i>	85,5	85,5	
Изучение тем и проблем, не выносимых на лекции и практические занятия	60	60	
Подготовка к экзамену	25,5	25,5	
Консультации и промежуточная аттестация	10,5	10,5	
Вид контроля (зачет, диф.зачет, экзамен)	-	экзамен	

5. Содержание дисциплины

№ раздела	Наименование разделов дисциплины	Объем аудиторных занятий по видам в часах			
		Всего	Л	ПЗ	ЛР
1	Введение в большие данные и Hadoop	14	8	6	0
2	Экосистема Hadoop	18	6	12	0
3	NoSQL-системы	16	10	6	0

5.1. Лекции

№ лекции	№ раздела	Наименование или краткое содержание лекционного занятия	Кол-во часов
1	1	Введение в большие данные и распределенные вычисления	2
2	1	Введение в платформу Hadoop	2
3	1	Распределенная файловая система Hadoop (HDFS)	2
4	1	Технология MapReduce	2
5	2	Экосистема Hadoop. Введение в Pig и СУБД Apache Hive	2
6	2	Введение в ETL. Брокер сообщений Apache Kafka	2
7	2	Apache Spark. Машинное обучение в Apache Spark	2
8	3	Классификация NoSQL-систем. NoSQL-системы "ключ-значение"	2
9	3	Документно-ориентированные NoSQL-системы	2
10	3	Колоночные NoSQL-системы	2
11	3	Графовые NoSQL-системы	2
12	3	Различия между SQL и NoSQL. Теорема CAP. Согласованность данных	2

5.2. Практические занятия, семинары

№ занятия	№ раздела	Наименование или краткое содержание практического занятия, семинара	Кол-во часов
-----------	-----------	---	--------------

1	1	Работа с HDFS	2
2	1	Разработка MapReduce-приложения	4
3	2	Работа с Apache Kafka	6
4	2	Разработка приложения в Apache Spark	6
5	3	Разработка приложения для анализа и визуализация данных, хранящихся в NoSQL-системе	6

5.3. Лабораторные работы

Не предусмотрены

5.4. Самостоятельная работа студента

Выполнение СРС			
Подвид СРС	Список литературы (с указанием разделов, глав, страниц) / ссылка на ресурс	Семестр	Кол-во часов
Изучение тем и проблем, не выносимых на лекции и практические занятия	[Доп. лит., 3] Гл. 1-11, с. 22-335; [Осн. лит., 4], с. 35-495	8	60
Подготовка к экзамену	[Осн. лит., 1], Часть 3, с. 331–400; [Осн. лит., 2], Гл. 24-25, с. 446-470; [Доп. лит., 5] с. 5-47	8	25,5

6. Фонд оценочных средств для проведения текущего контроля успеваемости, промежуточной аттестации

Контроль качества освоения образовательной программы осуществляется в соответствии с Положением о балльно-рейтинговой системе оценивания результатов учебной деятельности обучающихся.

6.1. Контрольные мероприятия (КМ)

№ КМ	Се-местр	Вид контроля	Название контрольного мероприятия	Вес	Макс. балл	Порядок начисления баллов	Учитывается в ПА
1	8	Текущий контроль	Минитест 1: большие данные	2	5	Минитест проводится в виде электронного теста в конце лекционного занятия. Тест содержит 5 вопросов, за каждый из которых можно получить максимум 1 балл. Студент получает 1 балл за вопрос, если ответ полностью верный, 0 баллов - иначе. Оценка студента за тест - это сумма баллов за каждый вопрос. Время, отведенное на опрос, 10 минут.	экзамен
2	8	Текущий контроль	Минитест 2: Hadoop	2	5	Минитест проводится в виде электронного теста в конце лекционного занятия. Тест содержит 5 вопросов, за каждый из которых можно получить максимум 1 балл. Студент получает 1 балл за вопрос,	экзамен

						если ответ полностью верный, 0 баллов - иначе. Оценка студента за тест - это сумма баллов за каждый вопрос. Время, отведенное на опрос, 10 минут.	
3	8	Текущий контроль	Минитест 3: HDFS	2	5	Минитест проводится в виде электронного теста в конце лекционного занятия. Тест содержит 5 вопросов, за каждый из которых можно получить максимум 1 балл. Студент получает 1 балл за вопрос, если ответ полностью верный, 0 баллов - иначе. Оценка студента за тест - это сумма баллов за каждый вопрос. Время, отведенное на опрос, 10 минут.	экзамен
4	8	Текущий контроль	Минитест 4: MapReduce	2	5	Минитест проводится в виде электронного теста в конце лекционного занятия. Тест содержит 5 вопросов, за каждый из которых можно получить максимум 1 балл. Студент получает 1 балл за вопрос, если ответ полностью верный, 0 баллов - иначе. Оценка студента за тест - это сумма баллов за каждый вопрос. Время, отведенное на опрос, 10 минут.	экзамен
5	8	Текущий контроль	Минитест 5: Pig и Hive	2	5	Минитест проводится в виде электронного теста в конце лекционного занятия. Тест содержит 5 вопросов, за каждый из которых можно получить максимум 1 балл. Студент получает 1 балл за вопрос, если ответ полностью верный, 0 баллов - иначе. Оценка студента за тест - это сумма баллов за каждый вопрос. Время, отведенное на опрос, 10 минут.	экзамен
6	8	Текущий контроль	Минитест 6: Apache Kafka	2	5	Минитест проводится в виде электронного теста в конце лекционного занятия. Тест содержит 5 вопросов, за каждый из которых можно получить максимум 1 балл. Студент получает 1 балл за вопрос, если ответ полностью верный, 0 баллов - иначе. Оценка студента за тест - это сумма баллов за каждый вопрос. Время, отведенное на опрос, 10 минут.	экзамен
7	8	Текущий контроль	Минитест 7: Apache Spark	2	5	Минитест проводится в виде электронного теста в конце лекционного занятия. Тест содержит 5 вопросов, за каждый из которых можно получить максимум 1 балл. Студент получает 1 балл за вопрос,	экзамен

						если ответ полностью верный, 0 баллов - иначе. Оценка студента за тест - это сумма баллов за каждый вопрос. Время, отведенное на опрос, 10 минут.	
8	8	Текущий контроль	Минитест 8: NoSQL-системы "ключ-значение"	2	5	Минитест проводится в виде электронного теста в конце лекционного занятия. Тест содержит 5 вопросов, за каждый из которых можно получить максимум 1 балл. Студент получает 1 балл за вопрос, если ответ полностью верный, 0 баллов - иначе. Оценка студента за тест - это сумма баллов за каждый вопрос. Время, отведенное на опрос, 10 минут.	экзамен
9	8	Текущий контроль	Минитест 9: Документо-ориентированные NoSQL-системы	2	5	Минитест проводится в виде электронного теста в конце лекционного занятия. Тест содержит 5 вопросов, за каждый из которых можно получить максимум 1 балл. Студент получает 1 балл за вопрос, если ответ полностью верный, 0 баллов - иначе. Оценка студента за тест - это сумма баллов за каждый вопрос. Время, отведенное на опрос, 10 минут.	экзамен
10	8	Текущий контроль	Минитест 10: колоночные NoSQL-системы	2	5	Минитест проводится в виде электронного теста в конце лекционного занятия. Тест содержит 5 вопросов, за каждый из которых можно получить максимум 1 балл. Студент получает 1 балл за вопрос, если ответ полностью верный, 0 баллов - иначе. Оценка студента за тест - это сумма баллов за каждый вопрос. Время, отведенное на опрос, 10 минут.	экзамен
11	8	Текущий контроль	Минитест 11: графовые NoSQL-системы	2	5	Минитест проводится в виде электронного теста в конце лекционного занятия. Тест содержит 5 вопросов, за каждый из которых можно получить максимум 1 балл. Студент получает 1 балл за вопрос, если ответ полностью верный, 0 баллов - иначе. Оценка студента за тест - это сумма баллов за каждый вопрос. Время, отведенное на опрос, 10 минут.	экзамен
12	8	Текущий контроль	Минитест 12: CAP	2	5	Минитест проводится в виде электронного теста в конце лекционного занятия. Тест содержит 5 вопросов, за каждый из которых можно получить максимум 1 балл. Студент получает 1 балл за вопрос,	экзамен

						если ответ полностью верный, 0 баллов - иначе. Оценка студента за тест - это сумма баллов за каждый вопрос. Время, отведенное на опрос, 10 минут.	
13	8	Текущий контроль	ПЗ 1. Работа с HDFS	12	1	1 балл: задание полностью выполнено 0 баллов: задание не выполнено	экзамен
14	8	Текущий контроль	ПЗ 2. Разработка MapReduce-приложения	16	1	1 балл: задание полностью выполнено 0 баллов: задание не выполнено	экзамен
15	8	Текущий контроль	ПЗ 3. Работа с Apache Kafka	16	3	1 балл: задание полностью выполнено 0 баллов: задание не выполнено	экзамен
16	8	Текущий контроль	ПЗ 4. Разработка приложения в Apache Spark	16	1	1 балл: задание полностью выполнено 0 баллов: задание не выполнено	экзамен
17	8	Текущий контроль	ПЗ 5. Разработка базы данных с использованием NoSQL-решения	16	1	1 балл: задание полностью выполнено 0 баллов: задание не выполнено	экзамен
18	8	Промежуточная аттестация	Итоговое тестирование	-	100	Промежуточная аттестация включает компьютерное тестирование. Контрольное мероприятие промежуточной аттестации проводится во время экзамена. Тест состоит из 20 случайных равноценных вопросов, позволяющих оценить сформированность компетенций. На ответы отводится 45 минут.	экзамен

6.2. Процедура проведения, критерии оценивания

Вид промежуточной аттестации	Процедура проведения	Критерии оценивания
экзамен	<p>При оценивании результатов учебной деятельности обучающегося по дисциплине используется балльно-рейтинговая система оценивания результатов учебной деятельности обучающихся (Положение о БРС утверждено приказом ректора от 24.05.2019 г. № 179, в редакции приказа ректора от 10.03.2022 г. № 25-13/09). Процедура прохождения промежуточной аттестации осуществляется согласно Положению о текущем контроле успеваемости и промежуточной аттестации (приказ ректора от 27.02.2024 № 33-13/09). Оценка за дисциплину формируется на основе полученных оценок за контрольно-рейтинговые мероприятия текущего контроля следующим образом: • Отлично: Величина рейтинга обучающегося по дисциплине 85...100 %. • Хорошо: Величина рейтинга обучающегося по дисциплине 75...84 %. • Удовлетворительно: Величина рейтинга обучающегося по дисциплине 60...74 %. • Неудовлетворительно: Величина рейтинга обучающегося по дисциплине 0...59 %. Если студент согласен с оценкой, полученной по результатам текущего контроля, то он может в день, предшествующий промежуточной аттестации дать свое согласие на автомат в личном кабинете. В случае явки студента на промежуточную</p>	В соответствии с пп. 2.5, 2.6 Положения

ПК-7	Имеет практический опыт: обработки и анализа данных в экосистеме Hadoop		+	+	+	+	+									+	+					+	+
------	---	--	---	---	---	---	---	--	--	--	--	--	--	--	--	---	---	--	--	--	--	---	---

Типовые контрольные задания по каждому мероприятию находятся в приложениях.

7. Учебно-методическое и информационное обеспечение дисциплины

Печатная учебно-методическая документация

а) основная литература:

Не предусмотрена

б) дополнительная литература:

Не предусмотрена

в) отечественные и зарубежные журналы по дисциплине, имеющиеся в библиотеке:

Не предусмотрены

г) методические указания для студентов по освоению дисциплины:

1. Методические указания для студентов и преподавателей по освоению и организации самостоятельной работы студентов

2. Методические указания для самостоятельной работы студента по практическим занятиям

из них: учебно-методическое обеспечение самостоятельной работы студента:

1. Методические указания для студентов и преподавателей по освоению и организации самостоятельной работы студентов

2. Методические указания для самостоятельной работы студента по практическим занятиям

Электронная учебно-методическая документация

№	Вид литературы	Наименование ресурса в электронной форме	Библиографическое описание
1	Основная литература	Электронно-библиотечная система издательства Лань	Григорьев, Ю. А. Реляционные базы данных и системы NoSQL : учебное пособие / Ю. А. Григорьев, А. Д. Плутенко, О. Ю. Плужникова. — Благовещенск : АмГУ, 2018. — 424 с. — ISBN 978-5-93493-308-2. — Текст : электронный // Лань : электронно-библиотечная система. — URL: https://e.lanbook.com/book/156492
2	Основная литература	Электронно-библиотечная система издательства Лань	Осипов, Д. Л. Технологии проектирования баз данных / Д. Л. Осипов. — Москва : ДМК Пресс, 2019. — 498 с. — ISBN 978-5-97060-737-4. — Текст : электронный // Лань : электронно-библиотечная система. — URL: https://e.lanbook.com/book/131692
3	Дополнительная литература	Электронно-библиотечная система издательства Лань	Маккинни, У. Python и анализ данных / У. Маккинни ; перевод с английского А. А. Слинкина. — 2-ое изд., испр. и доп. — Москва : ДМК Пресс, 2020. — 540 с. — ISBN 978-5-97060-590-5. — Текст : электронный // Лань : электронно-библиотечная система. — URL:

			https://e.lanbook.com/book/131721
4	Основная литература	Электронно-библиотечная система издательства Лань	Перрен, Ж. -. Spark в действии / Ж. -. Перрен ; перевод с английского А. В. Снастина. — Москва : ДМК Пресс, 2021. — 636 с. — ISBN 978-5-97060-879-1. — Текст : электронный // Лань : электронно-библиотечная система. — URL: https://e.lanbook.com/book/241001
5	Дополнительная литература	Электронно-библиотечная система издательства Лань	Бутаков, Н. А. Обработка больших данных с Apache Spark : учебно-методическое пособие / Н. А. Бутаков, М. В. Петров, Д. Насонов. — Санкт-Петербург : НИУ ИТМО, 2019. — 50 с. — Текст : электронный // Лань : электронно-библиотечная система. — URL: https://e.lanbook.com/book/136573

Перечень используемого программного обеспечения:

1. РСК Технологии-Система "Персональный виртуальный компьютер" (ПВК) (MS Windows, MS Office, открытое ПО)(бессрочно)

Перечень используемых профессиональных баз данных и информационных справочных систем:

Нет

8. Материально-техническое обеспечение дисциплины

Вид занятий	№ ауд.	Основное оборудование, стенды, макеты, компьютерная техника, предустановленное программное обеспечение, используемое для различных видов занятий
Лекции		Мультимедийный проектор
Экзамен		Компьютерный класс или ПВК-класс
Практические занятия и семинары		ПВК-класс