

<p>Документ подписан простой электронной подписью Информация о владельце: ФИО: Таскаев Сергей Валерьевич Должность: Ректор Дата подписания: 18.03.2025 14:53:16 Уникальный программный ключ: 04c19ed8bf98f7b6cb77a486b0a8788b832a7277</p>	 <p>МИНОБРАЗОВАНИЯ И НАУКИ РОССИИ Федеральное государственное бюджетное образовательное учреждение высшего образования «Челябинский государственный университет» (ФГБОУ ВО «ЧелГУ»)</p>	<p>Рабочая программа дисциплины "Администрирование и проектирование хранилищ больших данных" по направлению подготовки (специальности) 01.03.02 "Прикладная математика и информатика" направленности (профиль) Прикладная математика и искусственный интеллект ФГБОУ ВО «ЧелГУ»</p>	<p>стр. 1</p>
---	---	---	---------------

Рабочая программа дисциплины (модуля)*

Администрирование и проектирование хранилищ больших данных

Направление подготовки (специальность)

01.03.02 Прикладная математика и информатика

Направленность (профиль)

Прикладная математика и искусственный интеллект

Присваиваемая квалификация (степень)

бакалавр

Форма обучения

очная

Год(ы) набора 2024

*Рабочая программа дисциплины (модуля) адаптирована для инклюзивного обучения инвалидов и лиц с ограниченными возможностями здоровья

Челябинск 2024 г.



Содержание

1. Цели освоения дисциплины
2. Место дисциплины в структуре ОПОП
3. Компетенции обучающегося, формируемые в результате освоения дисциплины (модуля)
4. Объем дисциплины (модуля)
5. Структура и содержание дисциплины (модуля)
6. Фонд оценочных средств
 - 6.1. Перечень видов оценочных средств
 - 6.2. Типовые контрольные задания и иные материалы для текущей аттестации
 - 6.3. Типовые контрольные вопросы и задания для промежуточной аттестации
 - 6.4. Критерии оценивания
7. Учебно-методическое и информационное обеспечение дисциплины (модуля)
 - 7.1. Рекомендуемая литература
 - 7.2. Перечень ресурсов информационно-телекоммуникационной сети "Интернет"
 - 7.3. Перечень информационных технологий
8. Материально-техническое обеспечение дисциплины (модуля)
9. Методические указания для обучающихся по освоению дисциплины (модуля)
10. Специальные условия освоения дисциплины обучающимися с инвалидностью и ограниченными возможностями здоровья



1. ЦЕЛИ ОСВОЕНИЯ ДИСЦИПЛИНЫ

Цель курса: углубленное изучение и получение практических навыков программирования баз данных, изучение современных подходов к использованию и перспективных направлений развития СУБД. Задачи курса: - дать знания в области технологий баз данных, достаточные для профессиональной деятельности по разработке интеллектуальных систем; - сориентировать студентов во множестве современных СУБД и связанных с ними информационных технологий хранения и обработки больших массивов данных; - осветить теоретические и организационно-методические вопросы построения и функционирования систем, основанных на концепции баз данных, в том числе различные методологии моделирования и проектирования баз данных, обеспечения безопасности и надежности систем хранения и обработки больших данных; - дать практические навыки проектирования, администрирования и использования баз данных, а также разработки прикладного программного обеспечения.

2. МЕСТО ДИСЦИПЛИНЫ В СТРУКТУРЕ ОПОП

Цикл (раздел) ОПОП: Б1.В.1.08

2.1 Требования к предварительной подготовке обучающегося:

Изучение данной дисциплины опирается на знания, полученные при освоении следующих дисциплин:

Сбор, анализ и предобработка данных

2.2 Дисциплины и практики, для которых освоение данной дисциплины (модуля) необходимо как предшествующее:

Высокопроизводительные параллельные вычисления на кластерных системах

Преддипломная практика

Практика по получению профессиональных умений и опыта профессиональной деятельности

3. КОМПЕТЕНЦИИ ОБУЧАЮЩЕГОСЯ, ФОРМИРУЕМЫЕ В РЕЗУЛЬТАТЕ ОСВОЕНИЯ ДИСЦИПЛИНЫ (МОДУЛЯ)

В результате освоения дисциплины обучающийся должен

3.1	Знать:
3.1.1	ПК-10.1. 3-1. виды представления данных, методы поиска и парсинга данных;
3.1.2	ПК-10.1. 3-2. уровни представления данных (ODS, DDL, семантический слой, модель данных);
3.1.3	ПК-10.1. 3-3. основные инструменты и технологии Data Science
3.1.4	ПК-11.1. 3-1. общедоступные репозитории и специализированные библиотеки, содержащие наборы больших данных;
3.1.5	ПК-11.1. 3-2. принципы работы экосистемы Hadoop. фреймворка SPARK; [ПК-8.1. 3-3.] устройство интерфейсов между реляционными SQL хранилищами данных и нереляционными NoSQL-хранилищами данных;
3.1.6	ПК-11.1. 3-4. предметно-ориентированные языки;
3.1.7	ПК-11.2. 3- 2. устройство и принципы работы систем обработки и анализа больших массивов данных (SQL, NoSQL, Hadoop, ETL)
3.2	Уметь:
3.2.1	ПК-10.1. У-1. отделять достоверные источники данных от сомнительных, осуществлять критический отбор данных, проверять их на целостность и непротиворечивость
3.2.2	ПК-11.1. У-1. настраивать и оптимизировать конфигурацию программного и аппаратного обеспечения с целью интеграции больших данных;
3.2.3	ПК-11.1. У-2. разрабатывать программное обеспечение для очистки и валидации наборов больших данных;
3.2.4	ПК-11.1. У-3. выполнять потоковую обработку данных (data streaming, event processing);



Рабочая программа дисциплины "Администрирование и проектирование хранилищ больших данных" по направлению подготовки (специальности) 01.03.02 "Прикладная математика и информатика" направленности (профилю) Прикладная математика и искусственный интеллект ФГБОУ ВО «ЧелГУ»

стр. 4

3.2.5	ПК-11.1. У-5. использовать языки запросов, в том числе нереляционных, для поддержки различных типов данных (например, XML, RDF, JSON, мультимедиа) и операций с большими данными (например, матричные операции);
3.2.6	ПК-11.2. У-3. использовать системы обработки и анализа больших массивов данных (SQL, NoSQL, Hadoop, ETL, процессы и инструменты)
3.3	Владеть:
3.3.1	ПК-10.1. О-1. использования инструментов и библиотек для Data Science для поиска данных, технологии Data Science для подготовки и разметки структурированных и неструктурированных данных для машинного обучения
3.3.2	ПК-11.1. О-1. участия в разработке программного обеспечения для анализа больших данных

4. ОБЪЕМ ДИСЦИПЛИНЫ (МОДУЛЯ)

Общая трудоемкость	4 ЗЕТ
Часов по учебному плану : 144	Виды контроля в семестрах: экзамены 6 курсовые работы 6
в том числе :	
аудиторные занятия : 64	
самостоятельная работа : 68,5	
контактная работа: 75,5 ИКР: 11,5	

5. СТРУКТУРА И СОДЕРЖАНИЕ ДИСЦИПЛИНЫ (МОДУЛЯ)

Код занятия	Наименование разделов и тем /вид занятия/	Семестр / Курс	Часов	Литература
	Раздел 1. Концепции баз, банков, хранилищ данных.			
1.1	Эволюция систем хранения данных. Архитектура приложений для работы базами и хранилищами данных. Модель жизненного цикла баз данных. Трехуровневая модель разработки баз данных: концептуальное, логическое и физическое проектирование. Задачи каждого этапа. Концептуальная модель как модель предметной области, анализ требований к системе, построение и проверка ER-модели. Пример концептуального проектирования. /Лек/	6	2	Л1.2 Л1.3 Л1.4 Л1.1Л2.1
1.2	Переход от ER-модели к схеме реляционной базы данных. Детализация связей, определение доменов, ключей. Проверка условий нормализации и путей выполнения транзакций. Реализация базы данных в среде PostgreSQL Server: создание таблиц, пользовательских типов данных, ограничений целостности, представлений, функций пользователя и триггеров, выполнение транзакций. Пример логического и физического проектирования. /Лек/	6	4	Л1.2 Л1.3 Л1.4 Л1.1Л2.1
1.3	Основные задачи администратора базы данных: оптимизация производительности базы данных, обеспечение и контроль доступа к базе данных, роли и разграничение прав пользователей БД, обеспечение безопасности в базе данных, резервирование и восстановление базы данных, обеспечение целостности баз данных. /Лек/	6	2	Л1.2 Л1.3 Л1.4 Л1.1Л2.1
1.4	КР1. Концептуальное проектирование БД, Архитектура ANSI Spars. /Лаб/	6	4	Л1.2 Л1.3 Л1.4 Л1.1Л2.1
1.5	КР2. Логическое проектирование БД /Лаб/	6	4	Л1.2 Л1.3 Л1.4 Л1.1Л2.1
1.6	КР3. Реализация БД PostgreSQL. Создание БД, реализация ограничений, обеспечивающих целостность /Лаб/	6	4	Л1.2 Л1.3 Л1.4 Л1.1Л2.1
	Раздел 2. Типы современных СУБД			



2.1	Распределенные СУБД. Функции и архитектура РСУБД. Разработка распределенных баз данных, фрагментация, обеспечение прозрачности РСУБД, правила Дейта для РСУБД. Управление распределенными транзакциями, блокирующие протоколы, протоколы с временными отметками. Восстановление распределенных баз данных. Модель распределенной обработки транзакций X/Open. Серверы репликации /Лек/	6	2	Л1.2 Л1.3 Л1.4 Л1.1Л2.1
2.2	Объектно-ориентированные модели данных, манифест объектно-ориентированных СУБД (ООСУБД). Проектирование ООСУБД: связи и ссылочная целостность, проектирование поведения. Стандарт ODMG, язык описания объектов ODL, объектно-ориентированный язык запросов OQL. Объектно-реляционные СУБД (ОРСУБД). Манифест баз данных третьего поколения. ОРСУБД Postgres. Стандарт SQL3. Обработка и оптимизация запросов. Сравнительная характеристика ОРСУБД и ООСУБД /Лек/	6	2	Л1.2 Л1.3 Л1.4 Л1.1Л2.1
2.3	Проблемы больших данных. Недостатки реляционной модели. NoSQL решения, CAP теорема. Типы NoSQL систем, классификация: хранилища ключ-значение, документо-ориентированные, колоночные, хранилища на графах. Области применения. Примеры СУБД. /Лек/	6	4	Л1.2 Л1.3 Л1.4 Л1.1Л2.1
2.4	С1. Семинар. Современные базы данных /Лаб/	6	4	Л1.2 Л1.3 Л1.4 Л1.1Л2.1
2.5	ЛР1. MongoDB. Создание БД, создание, удаление и обновление документов, выборка данных из коллекции. /Лаб/	6	4	Л1.2 Л1.3 Л1.4 Л1.1Л2.1
Раздел 3. OLAP-технологии, хранилища больших данных				
3.1	Обзор. Что такое Business intelligence (BI), задачи, этапы: информационный поиск, трансформация и очистка данных (ETL, ELT), аналитическая обработка (OLAP), бизнес аналитика, бизнес отчетность. Хранилища данных (ХД). /Лек/	6	2	Л1.2 Л1.3 Л1.4 Л1.1Л2.1
3.2	Архитектура ХД. Проектирование ХД: подход Р.Кимбалла, Б.Инмона, DataVault, Ancor. Сравнительная характеристика, достоинства и недостатки. /Лек/	6	2	Л1.2 Л1.3 Л1.4 Л1.1Л2.1
3.3	Принципы построения корпоративных ХД: Классические, виртуальные, облачные ХД. Логическое моделирование ХД, OLAP-куб, схемы звезда и снежинка, операции. Реляционный OLAP, многомерный, гибридный. /Лек/	6	2	Л1.2 Л1.3 Л1.4 Л1.1Л2.1
3.4	Процессы ETL, пакетная и потоковая обработка, достоинства и недостатки. Примеры реализации на Python. Обзор готовых инструментов интеграции данных. Отечественные ХД, обзор. /Лек/	6	2	Л1.2 Л1.3 Л1.4 Л1.1Л2.1
3.5	ЛР2. OLAP. Проектирование реляционного хранилища данных /Лаб/	6	4	Л1.2 Л1.3 Л1.4 Л1.1Л2.1
Раздел 4. Технологии работы с хранилищами больших данных				
4.1	Современные технологии работы с большими данными. Архитектура фреймворка Hadoop. Модули Hadoop Common, HDFS, YARN, Hadoop MapReduce. /Лек/	6	4	Л1.2 Л1.3 Л1.4 Л1.1Л2.1
4.2	Экосистема Apache Spark. Характеристика пакетов Spark Core, Spark SQL, Spark Streaming, MLLib, GraphX. Сравнение с Hadoop. Другие Open-Source системы, обзор: Apache Kafka, Scala, Apache Beam. /Лек/	6	4	Л1.2 Л1.3 Л1.4 Л1.1Л2.1
4.3	ЛР3. Hadoop. Разработка распределенных программ /Лаб/	6	4	Л1.2 Л1.3 Л1.4 Л1.1Л2.1
4.4	ЛР4. Spark. Распределенная обработка больших данных /Лаб/	6	4	Л1.2 Л1.3 Л1.4 Л1.1Л2.1
Раздел 5. Самостоятельная работа				
5.1	Выполнение и защита курсовой работы /Ср/	6	40,5	Л1.2 Л1.3 Л1.4 Л1.1Л2.1



Рабочая программа дисциплины "Администрирование и проектирование хранилищ больших данных" по направлению подготовки (специальности) 01.03.02 "Прикладная математика и информатика" направленности (профилю) Прикладная математика и искусственный интеллект ФГБОУ ВО «ЧелГУ»

стр. 6

5.2	Подготовка к экзамену /Ср/	6	14	Л1.2 Л1.3 Л1.4 Л1.1Л2.1
5.3	Проработка материала лекций, подготовка к лабораторным работам /Ср/	6	8	Л1.2 Л1.3 Л1.4 Л1.1Л2.1
5.4	Подготовка к семинару /Ср/	6	6	Л1.2 Л1.3 Л1.4 Л1.1Л2.1
Раздел 6. Иная контактная работа				
6.1	Индивидуальные консультации, текущий контроль /ИКР /ИКР/	6	11,5	Л1.2 Л1.3 Л1.4 Л1.1Л2.1

6. ФОНД ОЦЕНОЧНЫХ СРЕДСТВ

6.1. Перечень видов оценочных средств

Семинар;
Лабораторная работа 1-4;
Защита курсовой работы;
Экзамен

6.2. Типовые контрольные задания и иные материалы для текущей аттестации

Примерные варианты мероприятий текущего контроля приведены в приложении.

6.3. Типовые контрольные вопросы и задания для промежуточной аттестации

Вопросы к экзамену по курсу «Администрирование и проектирование хранилищ больших данных»

1. Концепции баз, банков, хранилищ данных
 - 1.1. Жизненный цикл БД
 - 1.2. Постановка задачи разработки, инфологическая модель
 - 1.3. Концептуальное проектирование и ER-модель
 - 1.4. Логическое проектирование. Нормализация
 - 1.5. Безопасность данных в PostgreSQL
 - 1.6. Работа с транзакциями в PostgreSQL. Блокировки. Уровни изоляции транзакций
 - 1.7. Использование индексов в PostgreSQL
 - 1.8. Оптимизация запросов для достижения максимальной производительности
 - 1.9. Резервное копирование и восстановление в PostgreSQL
 - 1.10. Резервное копирование и репликация в PostgreSQL
2. Типы современных СУБД
 - 2.1. Типы современных СУБД
 - 2.2. CAP-теорема
 - 2.3. СУБД MongoDB. Особенности. Область применения.
 - 2.4. Клиент-серверные СУБД
 - 2.5. Модели распределения функций
 - 2.6. Распределенные БД. Правила Дейта. Фрагментация и репликация
 - 2.7. Распределенные транзакции
 - 2.8. Преимущества и недостатки распределенных СУБД
 - 2.9. Документо-ориентированные БД
 - 2.10. Подход NoSQL
 - 2.11. БД ключ-значение
 - 2.12. New SQL БД
3. OLAP-технологии, хранилища больших данных
 - 3.1. Модель сервера баз данных. Достоинства и недостатки.
 - 3.2. Модель сервера приложений. Достоинства и недостатки
 - 3.3. Архитектура хранилищ данных
 - 3.4. OLAP-куб. Схемы звезда и снежинка.
 - 3.5. Требования к OLAP-инструментам
4. Технологии работы с хранилищами больших данных
 - 4.1. Что такое «большие данные»
 - 4.2. Принципы работы с большими данными
 - 4.3. Архитектура Apache Hadoop
 - 4.4. Архитектура Spark
 - 4.5. Распределенная обработка MapRaduce



6.4. Критерии оценивания

Экзамен выставляется, в соответствии с БРС, на основе результатов текущего контроля. Студент может повысить свой рейтинг, пройдя контрольное мероприятие промежуточной аттестации.

Итоговая оценка выставляется по 100-балльной шкале (в процентном соотношении от общей суммы баллов), исходя из полученной суммы баллов:

- От 0 до 50 баллов – «неудовлетворительно»
- От 51 до 65 баллов – «удовлетворительно»
- От 66 до 80 баллов – «хорошо»
- От 81 балла – «отлично».

Экзамен проводится в очной форме по экзаменационным билетам. Процедура прохождения зачета не является обязательной если по результатам текущего контроля БРС у студента положительная оценка и он с ней согласен. В каждом билете 3 теоретических вопроса. Экзамен принимается в устной форме. Студент должен находиться в аудитории на протяжении всей процедуры экзамена. Число студентов, одновременно находящихся в аудитории, где сдается экзамен, не более 8 человек. На подготовку к ответу студенту отводится не более 30 мин. Когда обучающийся будет готов к ответу, ему задаются контрольные вопросы по содержанию билета. Студент должен УСТНО ответить на эти вопросы в течение 5 мин. На этом основании преподаватель выставляет оценку за экзамен. Оценка за курс в целом выставляется согласно БРС. Добор баллов осуществляется посредством выполнения заданий КМ4-КМ8.

Процедура защиты курсового проекта (КП) является обязательной. Защита КП проходит в открытой форме. На защите могут присутствовать студенты и преподаватели. Оценка по результатам защиты выставляется комиссией из преподавателей, назначаемых заведующим кафедрой. Для защиты курсового проекта студентом должны быть представлены: а) программа в виде исходного кода и выполняемого файла (архив необходимо заранее загрузить в электронный курс); б) отчет в электронном и печатном виде (файл с отчетом также должен быть загружен в электронный курс). Защита включает в себя доклад студента с демонстрацией отчета по КП перед комиссией в течении 7 минут, и последующие ответы на вопросы (например, «почему было реализовано именно таким образом», «имело ли смысл предусмотреть в программе такие-то функции» и т.п.);

Оценка суммируется из следующих оценок: 1) отчет содержит все необходимые разделы - 1 балл 2) нет ошибок в оформлении, речевых и грамматических ошибок- 1 балл 3) В устном изложении студент показывает знания теории БД, свободно оперирует терминами применительно к рассматриваемой задаче - 2 балла; имеются ошибки в терминологии - 1 балл; не владение понятиями и методами теории баз данных - 0 баллов 4) Студент может правильно и полно ответить на вопросы членов комиссии - 2 балла; частично правильные, не полные ответы - 1 балл; неверные ответы - 0 баллов

Оценка за курсовую работу суммируется из следующих оценок:

КР1. Концептуальное проектирование (максимально оценивается в 8 баллов)

1) сформулирована тема, содержательно описана предметная область системы - 1 балл; 2) указаны пользователи и все функциональные требования (автоматизируемые задачи пользователей) - 1 балл; 3) описан основной бизнес-процесс системы - 1 балл; 4) описан интерфейс (пункты меню, кнопки, используемые устройства взаимодействия) - 1 балл; 5) выявлены ограничения целостности данных: уникальные поля, функциональные зависимости и т.п., а также, если нужно, пользовательские ограничения - 1 балл; 6) ER-диаграмма предметной области: не содержит ошибок - 2 балла; содержит незначительные ошибки - 1 балл; построена не верно - 0 баллов 7) выполнена проверка возможности реализации основного бизнес-процесса. - 1 балл.

КР2. Логическое проектирование (максимально оценивается в 5 баллов)

1) ER-диаграмма преобразована в реляционную схему БД: правильно - 2 балла; имеются ошибки преобразования - 1 балл; схема БД имеет серьезные ошибки - 0 баллов 2) выделены объекты реализации целостности БД: правильно - 2 балла; выделены не все объекты - 1 балл; схема БД имеет серьезные ошибки - 0 баллов 3) проверено соответствие нормальным формам - 1 балл.

КР3. Реализация БД PostgreSQL (максимально оценивается в 4 балла)

1) созданы таблицы, индексы, триггеры, функции в БД - 2 балла; созданы не все объекты, обеспечивающие безопасность БД - 1 балл; имеются существенные ошибки - 0 баллов 2) настроены связи между таблицами, в т.ч. политики удаления и обновления - 1 балл 3) созданы роли пользователей БД, выполнено разграничение прав - 1 балл



Семинар (максимально оценивается в 6 баллов)

Проблемы современных хранилищ данных, решения 1) раскрытие темы доклада 2) четкость изложения, 2) соблюдение регламента 3) ответы на вопросы Полное раскрытие темы доклада, выводы логичны – 2 балла; тема раскрыта не полностью или отсутствуют выводы – 1 балл; тема не раскрыта – 0 баллов. Четкость и последовательность изложения – 1 балл, иначе – 0 баллов Соблюдение регламента – 1 балл, иначе – 0 баллов Содержательные ответы на вопросы слушателей – 2 балла: неуверенные ответы – 1 балл; неудовлетворительные ответы – 0 баллов

Лабораторные работы 1-4 (максимально оценивается в 5 баллов)

1. Задание выполнено полностью. Созданы все необходимые модули – 1 балл; созданы не все модули – 0 баллов; 2. Приложение отлажено и работоспособно -2 балла; имеются ошибки или реализованы не все функции – 1 балл; приложение не работоспособно – 0 баллов 3. Отчет оформлен в соответствии с требованиями – 1 балл, иначе – 0 баллов 4. Полные и верные ответы на контрольные вопросы – 1 балл, иначе – 0 баллов

7. УЧЕБНО-МЕТОДИЧЕСКОЕ И ИНФОРМАЦИОННОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ (МОДУЛЯ)

7.1. Рекомендуемая литература

7.1.1. Основная литература

	Авторы,	Заглавие	Издательство,	Ресурс
ЛП.1	Перрен Ж. -.	Spark в действии (https://e.lanbook.com/book/241001)	Москва : ДМК Пресс, 2021	ЭБС
ЛП.2	Волчихин В. И., Иванов А. И., Назаров И. Г., Фунтиков В. А., Язов Ю. К.	Нейросетевая защита персональных биометрических данных	Москва: Радиотехника, 2012	
ЛП.3	Дейт К. Дж., Птицын К. А.	Введение в системы баз данных	Москва : Вильямс, 2006	
ЛП.4	Осипов Д. Л.	Технологии проектирования баз данных (https://e.lanbook.com/book/131692)	Москва : ДМК Пресс, 2019	ЭБС

7.1.2. Дополнительная литература

	Авторы,	Заглавие	Издательство,	Ресурс
Л2.1	Шёниг Г. -.	PostgreSQL 11. Мастерство разработки (https://e.lanbook.com/book/131714)	Москва : ДМК Пресс, 2020	ЭБС

7.3 Перечень информационных технологий

7.3.1 Программное обеспечение

LMS Moodle

Python

PostgreSQL

OpenOffice

7.3.2 Профессиональные базы данных и информационно-справочные системы

Электронный каталог научной библиотеки ЧелГУ [Электронный ресурс] : база данных / Челябин. гос. ун-т. – Челябинск, 1992

eLIBRARY.RU [Электронный ресурс] : научная электронная библиотека [научной периодики на русском языке]. — Москва, [1999-]. - Доступ к полным текстам после регистрации из сети ЧелГУ. – URL: <http://elibrary.ru/defaultx.asp>.

Moodle [Электронный ресурс]: система дистанционного обучения : [база данных] / Челябин. гос. ун-т. – Челябинск, [б.г.]. – Доступ из сети ЧелГУ или, после регистрации из сети ун-та, из любой точки, имеющей доступ в интернет. – URL: <http://moodle.uio.csu.ru/login/index.php>.

Научная библиотека Челябинского государственного университета [Электронный ресурс] : [сайт] / Челябин. гос. ун-т. – Челябинск, [2001-]. – Режим доступа: <http://www.lib.csu.ru/>, свободный. – Загл. с экрана.

8. МАТЕРИАЛЬНО-ТЕХНИЧЕСКОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ (МОДУЛЯ)



Для реализации дисциплины используются учебные аудитории для проведения занятий лекционного типа, занятий семинарского типа, лабораторных занятий, групповых и индивидуальных консультаций, текущего контроля и промежуточной аттестации, а также помещения для самостоятельной работы.

Учебные аудитории укомплектованы специализированной мебелью и техническими средствами обучения (мультимедийное устройство, проектор, ноутбук или стационарный компьютер).

Помещения для самостоятельной работы обучающихся оснащены компьютерной техникой с подключением к сети "Интернет" и обеспечением доступа в электронную информационно-образовательную среду университета.

9. МЕТОДИЧЕСКИЕ УКАЗАНИЯ ДЛЯ ОБУЧАЮЩИХСЯ ПО ОСВОЕНИЮ ДИСЦИПЛИНЫ (МОДУЛЯ)

Учебный курс строится таким образом, чтобы способствовать созданию у обучающегося понятийно–теоретической базы, развитию умения практического решения задач, умения работать со справочной литературой.

Для успешного усвоения материала студенту необходимо получить достаточное количество баллов по следующим формам обучения:

1. Лекционная форма, которая предполагает посещение лекций.
2. Практическая форма занятий предполагает выполнение лабораторных работ, использование справочной литературы.
3. Самостоятельная форма работы предполагает изучение теоретических вопросов, выполнение практических заданий. Для их выполнения обучающемуся необходимо использование и изучение литературы по заданной теме.

В каждом семестре студенты выполняют цикл лабораторных работ, которые служат основой для допуска к экзамену. По окончании шестого семестра студенты сдают экзамен. Сдача экзамена предполагает выполнение теста. В тест входят вопросы и задачи.

Примеры заданий, образец итоговых задач, образцы вопросов теста к зачету прилагаются в разделе ФОС.

При обучении лиц с ограниченными возможностями здоровья электронное обучение, дистанционные образовательные технологии предусматривают возможность приема-передачи информации в доступных для них формах.

Реализация дисциплины с применением электронного обучения, дистанционных образовательных технологий (далее – ЭО, ДОТ) осуществляется на основании «Положения о реализации основных и дополнительных образовательных программ с применением электронного обучения и дистанционных образовательных технологий в федеральном государственном бюджетном образовательном учреждении высшего образования «Челябинский государственный университет», «Положения о порядке зачета обучающимися по основным профессиональным образовательным программам высшего образования в ФГБОУ ВО «ЧелГУ» результатов освоения в организациях, осуществляющих образовательную деятельность, учебных предметов, курсов, дисциплин (модулей), практик, дополнительных образовательных программ» посредством электронной информационно-образовательной среды ФГБОУ ВО «ЧелГУ». В исключительных случаях (форс-мажор и т.п.) при реализации образовательной деятельности с применением ЭО, ДОТ могут применять компоненты, не входящие в перечень электронной информационно-

10. СПЕЦИАЛЬНЫЕ УСЛОВИЯ ОСВОЕНИЯ ДИСЦИПЛИНЫ ОБУЧАЮЩИМИСЯ С ИНВАЛИДНОСТЬЮ И ОГРАНИЧЕННЫМИ ВОЗМОЖНОСТЯМИ ЗДОРОВЬЯ

Освоение дисциплины инвалидами и лицами с ограниченными возможностями здоровья осуществляется с использованием специальных технических средств и голо информационных технологий, предоставляемых Ресурсным учебно-методическим центром по обучению инвалидов и лиц с ограниченными возможностями здоровья ЧелГУ по запросу обучающегося.

1. Мобильные специальные технические средства для лиц с нарушениями зрения: портативный компьютер с вводом/выводом шрифтом Брайля с синтезатором речи «EIBraile-W14J G2»; ноутбуки с программной экранного доступа NVDA; электронные увеличители для удаленного просмотра; видеоувеличители портативные; тифлоплеер; цифровые диктофоны.

2. Мобильные специальные технические средства для лиц с нарушениями слуха: система свободного звукового поля со встроенной совместимостью с FM-устройствами; радиоклассы «Сонет-PCM» с передатчиком, заушным индуктором и индукционной петлей; система информационная для слабослышащих переносная «Исток» А2 со встроенным плеером – звуковым информатором; документ-камера; программируемые слуховые аппараты индивидуального пользования.

3. Ассистивные информационные технологии: программное обеспечение экранного доступа с синтезом речи NVDA; программы экранного увеличения; программы речевого синтеза для компьютеров и ноутбуков; программы речевого синтеза для мобильных устройств; экранная клавиатура; экранная лупа.

При необходимости для обучающихся с нарушениями зрения на рабочих местах для проведения практических или



лабораторных занятий устанавливается специальное программное обеспечение (программа речевой навигации NVDA, речевые синтезаторы, экранные лупы).

В учебные аудитории обеспечивается беспрепятственный доступ для обучающихся инвалидов и обучающихся с ограниченными возможностями здоровья. В каждой аудитории, где обучаются инвалиды и лица с ограниченными возможностями здоровья, предусматривается соответствующее количество мест для обучающихся с учетом нарушений их здоровья.

Для освоения дисциплины инвалидам и лицам с ограниченными возможностями здоровья предоставляется доступ к печатным источникам, имеющимся в научной библиотеке ЧелГУ, с помощью специальных технических средств; доступ к электронным источникам, представленным в форме электронного документа в фонде научной библиотеки ЧелГУ или электронно-библиотечных системах, с помощью специальных технических и программных средств (рабочее место для незрячего пользователя с программным обеспечением экранного доступа с синтезом речи NVDA, рабочее место с компьютерным роллером и клавиатурой CleVu с большими кнопками и с разделяющей клавиши накладкой).

Учебно-методические материалы для обучающихся из числа инвалидов и лиц с ограниченными возможностями здоровья предоставляются в формах, адаптированных к ограничениям их здоровья и восприятия информации:

Для лиц с нарушениями зрения:

- в печатной форме увеличенным шрифтом,
- в форме электронного документа,
- в форме аудиофайла,
- в печатной форме шрифтом Брайля.

Для лиц с нарушениями слуха:

- в печатной форме,
- в форме электронного документа.

Для лиц с нарушениями опорно-двигательного аппарата:

- в печатной форме,
- в форме электронного документа,
- в форме аудиофайла.

Данный перечень может быть конкретизирован в зависимости от контингента обучающихся.

Для инвалидов и лиц с ограниченными возможностями здоровья освоение дисциплины может быть частично или полностью осуществлено с использованием дистанционных образовательных технологий (Moodle, Adobe Connect Pro и пр.).

В освоении дисциплины инвалидами и лицами с ограниченными возможностями здоровья используется индивидуальная работа. Под индивидуальной работой подразумевается две формы взаимодействия с преподавателем: индивидуальная учебная работа (консультации), т.е. дополнительное разъяснение учебного материала и углубленное изучение материала с теми обучающимися, которые в этом заинтересованы, и индивидуальная воспитательная работа. Индивидуальные консультации направлены на индивидуализацию обучения и установлению воспитательного контакта между преподавателем и обучающимся инвалидом или обучающимся с ограниченными возможностями здоровья.

При проведении процедуры оценивания результатов обучения инвалидов и лиц с ограниченными возможностями здоровья по дисциплине обеспечивается выполнение следующих дополнительных требований в зависимости от индивидуальных особенностей, обучающихся:

- а) инструкция по порядку проведения процедуры оценивания предоставляется в доступной форме (устно, в письменной форме, в письменной форме шрифтом Брайля, устно с использованием услуг сурдопереводчика);
- б) доступная форма предоставления заданий оценочных средств (в печатной форме, в печатной форме увеличенным шрифтом, в печатной форме шрифтом Брайля, в форме электронного документа, задания зачитываются ассистентом, задания предоставляются с использованием сурдоперевода);
- в) доступная форма предоставления ответов на задания (письменно на бумаге, набор ответов на компьютере, письменно шрифтом Брайля, с использованием услуг ассистента, устно).

При проведении процедуры оценивания результатов обучения инвалидов и лиц с ограниченными возможностями здоровья предусматривается использование технических средств, необходимых им в связи с их индивидуальными особенностями. Эти средства могут быть предоставлены ЧелГУ или могут использоваться собственные технические средства. При необходимости инвалидам и лицам с ограниченными возможностями здоровья предоставляется дополнительное время для подготовки ответа на задания, процедура оценивания результатов обучения по дисциплине может проводиться в несколько этапов.

Проведение процедуры оценивания результатов обучения инвалидов и лиц с ограниченными возможностями здоровья допускается с использованием дистанционных образовательных технологий.

ООП «Прикладная математика и искусственный интеллект» по направлению
01.03.02 – Прикладная математика и информатика

Дисциплина «Администрирование и проектирование хранилищ больших
данных»

№ КМ	Вид КМ	Наименование КМ	Оценочные средства
1.	Курсовая работа/проект	КР1. Концептуальное проектирование	<p>Цель: получить навыки моделирования предметной области баз данных (БД)</p> <p>Ожидаемый результат: оформление технического задания на разработку базы данных</p> <p>Задание на КР:</p> <ol style="list-style-type: none"> 1) Выбрать или предложить свой вариант предметной области 2) Выполнить анализ предметной области для проектирования бд, описать инфологическую модель предметной области, пользователей системы и функции, которые система предоставляет пользователям 3) Построить концептуальную модель (ER-диаграмму) предметной области 4) Определить ограничения целостности БД, накладываемые предметной областью 5) Оформить отчет о работе в виде технического задания на разработку БД <p>Пример индивидуального задания</p> <p>1) Система для продажи авиабилетов</p> <p>Объекты: авиакомпании, аэропорты, типы самолетов (мест), экземпляры самолетов, расписание, пассажиры (Ф.И.О, паспорт, № билета), цены.</p> <p>Основная бизнес-функция: продать билет.</p> <p>Отчеты: списки пассажиров по каждому рейсу, расписание рейсов.</p> <p>Контрольные вопросы</p> <ol style="list-style-type: none"> 1) Перечислите основные этапы проектирования БД? 2) Определите соотношение понятия “сущность”, “связь”? 3) Последовательность концептуального проектирования для конкретной предметной области? 4) Правила обозначения атрибутов на ER-модели 5) Допустимы ли в ER-модели связи «многие-ко-многим» 6) Как мы проверяем непротиворечивость ER-модели?
2.	Курсовая работа/проект	КР2. Логическое проектирование	<p>Цель: получить навыки проектирования реляционных баз данных (БД)</p> <p>Ожидаемый результат: построение схемы реляционной базы данных для выбранной предметной области</p> <p>Задание:</p>

№ КМ	Вид КМ	Наименование КМ	Оценочные средства
	проект		<p>1) Преобразовать ER-модель, созданную в КР 1 в схему реляционной БД</p> <p>2) Определить требования к объектам реляционной модели</p> <p>3) Оформить отчет о работе</p> <p>Контрольные вопросы</p> <p>1) Каковы задачи, решаемые на этапе логического проектирования?</p> <p>2) Каковы базовые свойства реляционной модели данных?</p> <p>3) В чем состоят требования целостной части реляционной модели данных?</p> <p>4) Каковы общие свойства нормальных форм?</p> <p>5) Каковы условия нахождения отношений в первой нормальной форме?</p> <p>6) Каковы условия нахождения отношений во второй нормальной форме?</p> <p>7) Каковы условия нахождения отношений в третьей нормальной форме?</p> <p>8) Каковы условия нахождения отношений в третьей усиленной нормальной форме?</p> <p>9) В чем состоят общие требования обеспечения ограничений целостности?</p> <p>10) Каковы средства задания ограничений целостности в реляционной модели?</p>
3.	Курсовая работа/проект	КР3. Реализация БД PostgreSQL	<p>Цель: получить навыки создания БД по модели, разработанной в КР 2, на СУБД PostgreSQL</p> <p>Ожидаемый результат: Физическая реализация БД</p> <p>Порядок выполнения работы:</p> <p>1) В СУБД PostgreSQL Server (см. Документация к postgresql) создать все необходимые объекты реляционной БД: таблицы, связи между ними,</p> <p>2) Реализовать пользовательские ограничения предметной области: проверочные ограничения, индексы, триггеры</p> <p>3) Реализовать элементы программирования: функции и хранимые процедуры.</p> <p>Контрольные вопросы</p> <p>1) Что такое независимость, безопасность, целостность, защита данных?</p> <p>2) Когда нужно применять индексирование полей БД?</p> <p>3) Для чего нужен триггер?</p> <p>4) Какие типы триггеров PostgreSQL вы знаете?</p> <p>5) Как обеспечиваются целостность и независимость данных?</p> <p>6) Какими средствами можно задать права пользователей БД?</p> <p>7) Каковы недостатки реляционных БД?</p>
4.	Текущий контроль	Семинар; Проблемы современных хранилищ данных, решения	<p>Обсуждаемые вопросы</p> <ul style="list-style-type: none"> - обработка «грязных» данных; - производительность и масштабируемость - выбор источников данных для загрузки в хранилище данных <p>Примерные темы докладов</p> <ol style="list-style-type: none"> 1. Проблемы классических реляционных СУБД. Решения NoSQL 2. Требования к современным хранилищам данных (ХД). Типы ХД 3. Отечественные облачные ХД. 4. Резидентные базы данных. Где применяются.

№ КМ	Вид КМ	Наименование КМ	Оценочные средства
			<p>5. СУБД для хранилищ данных. Современные решения</p> <p>6. СУБД PostgreSQL</p> <p>7. СУБД Riak</p> <p>8. СУБД Hbase</p> <p>9. СУБД MongoDB</p> <p>10. СУБД CouchDB</p> <p>11. СУБД Neo4j</p> <p>12. СУБД Redis</p> <p>Контрольные вопросы</p> <ol style="list-style-type: none"> 1. Что входит в состав хранилища данных? 2. Каковы основные характеристики хранилища данных? 3. В чем разница между базой данных и хранилищем данных? 4. В чем разница между ETL и ELT? 5. В каком виде представляется информация в хранилищах данных? 6. Основные механизмы доступа к БД в сети Интернет. 7. Системы оперативной обработки информации – OLTP. Архитектура и назначение. 8. Системы консолидации и аналитической обработки информации – ELT. Архитектура и назначение. 9. Концепция многомерного представления данных – гиперкубы. 10. Измерения и факты в гиперкубах. Правила выбора измерений и фактов. 11. Формализация многомерного представления данных: метки, иерархии, ячейки, меры. 12. Операции над данными в гиперкубах: вращение, сечение (срез), свертка и детализация. 13. Агрегация в гиперкубах: виды агрегации. 14. Агрегация в гиперкубах – оценка числа агрегатов для двумерного случая. 15. Классификация архитектуры хранилищ данных. 16. Многомерные хранилища данных. 17. Различие концепций ХД и особенности построения.
5.	Текущий контроль	ЛР1. MongoDB	<p>Цель</p> <ol style="list-style-type: none"> а. Ознакомиться с нереляционной документоориентированной СУБД MongoDB б. Получить навыки работы с БД из среды Python. в. Получить навыки ETL трансформации данных при переносе их из одной БД в другую. <p>Задание</p> <p>Работа выполняется с данными, загруженными в PostgreSQL БД, созданную на занятии ПКЗ, для вашей предметной области.</p> <ol style="list-style-type: none"> 1. Проанализировать вашу предметную область и определить, какие данные Вы загружаете в MongoDB 2. В БД PostgreSQL написать запросы на выборку необходимых данных. 3. На Python написать программу: <ol style="list-style-type: none"> а. создать подключение к PostgreSQL; б. выполнить выборку необходимых данных, преобразование их в формат JSON. в. создать подключение к СУБД MongoDB, г. создать коллекцию документов, выполняя запросы к PostgreSQL

№ КМ	Вид КМ	Наименование КМ	Оценочные средства
			<p>д. загрузить эти документы в MongoDB, добавить в коллекцию не менее четырех различных документов.</p> <p>4. Вывести список документов в коллекции со всеми атрибутами, выполнив запрос к базе данных.</p> <p>5. Вывести основные атрибуты части документов коллекции, удовлетворяющих некоторому условию (использовать условия «меньше», «больше»)</p> <p>Контрольные вопросы</p> <ol style="list-style-type: none"> 1. Что является отличительной особенностью NoSQL БД? 2. В каком случае стоит применять NoSQL хранилища? 3. Что, согласно теореме CAP, возможно обеспечить в любой реализации распределённых вычислений? 4. Какое свойство означает, что транзакции не нарушают согласованность данных, то есть они переводят базу данных из одного корректного состояния в другое? 5. Какой способ хранения данных используется в MongoDB? 6. Что относится к плюсам репликации? 7. Что относится к преимуществам нереляционных БД? 8. На какие три группы подразделяют пользователей в MongoDB? 9. Какие три свойства фигурируют в определении теоремы CAP?
6.	Текущий контроль	ЛР2. OLAP. Проектирование реляционного хранилища данных	<p>Цель: Ознакомиться с технологией и архитектурой хранилищ данных.</p> <p>Ожидаемые результаты: Получить навыки построения многомерных моделей данных и способы их обработки средствами Excel</p> <p>Порядок выполнения работы:</p> <ol style="list-style-type: none"> 1) постройте модель для анализа в виде «снежинка», используя вашу предметную область. Можно, например, проанализировать объемы продаж за различные временные интервалы; 2) напишите скрипты для создания OLAP-куба из вашей базы данных; 3) выгрузите OLAP-куб в Excel, как показано в рекомендациях; 4) постройте график или диаграмму временной зависимости для анализируемых данных; 5) оформите отчет о работе; <p>Контрольные вопросы</p> <ol style="list-style-type: none"> 1. Каковы основные различия между OLAP и OLTP системами? 2. Где используется OLAP? 3. Какая модель данных лежит в основе OLAP технологии? 4. Какие схемы используются при построении многомерной модели данных OLAP? 5. Сущность и назначение операции разрезания (slice) куба OLAP 6. Сущность и назначение иерархий значений в измерениях куба OLAP 7. Сущность и назначение Хранилищ данных 8. Основные достоинства хранения информации в многомерном кубе. 9. Дайте характеристику основным понятиям многомерной модели: Измерение, Показатель.

№ КМ	Вид КМ	Наименование КМ	Оценочные средства
			<p>10. Объясните назначение операций манипулирования с Измерениями: срез, вращение, детализация и агрегация.</p> <p>11. Назначение и содержание иерархии отношений в многомерном кубе.</p> <p>12. Особенности работы пользователей с OLAP-системами.</p> <p>13. Структура хранилища данных по схеме «звезда».</p> <p>14. Классификация OLAP-продуктов по способу хранения данных: MOLAP, ROLAP и HOLAP.</p>
7.	Текущий контроль	ЛР 3 Hadoop	<p>Цель: Ознакомиться с технологией распределенных вычислений и архитектурой фреймворка Hadoop.</p> <p>Задание Напишите программу на Python, используя фреймворк Hadoop, согласно вашему варианту для датасета “brooklyn_sales_map.csv”.</p> <p>Пример индивидуального задания Вариант 1.</p> <ul style="list-style-type: none"> – Найдите среднюю стоимость жилья (sale_price) и выведите новую таблицу, содержащую две колонки – стоимость жилья и отклонение стоимости от среднего значения. – Найдите среднюю стоимость жилья (sale_price) для каждого района. – Выведите среднюю полную площадь жилья (gross_sqft) для всех сочетаний налоговых категорий (tax_class) и лет продажи (year_of_sale). – Выведите таблицу, содержащую количество пустых (null) значений для каждой колонки. <p>Контрольные вопросы</p> <ol style="list-style-type: none"> 1. Что такое большие данные? 2. Объясните пять V больших данных 3. Сколько форматов ввода существует в Hadoop? 4. Что такое Thread? 5. Кто использует Hadoop? 6. Каковы основные функции Hadoop? 7. В каких трех режимах может работать Hadoop? 8. В каких операционных системах работает Hadoop? 9. Что делает JobTracker? 10. Расскажите о различных методах редуктора 11. Что делает команда jps? 12. Как вы проверяете NameNode при использовании команды jps? 13. В чем разница между левым полусоединением и внутренним соединением? 14. Как вы развертываете решение для работы с большими данными? 15. Каковы компоненты HDFS? 16. Какие минусы или опасные места HDFS? 17. Что входит в экосистему Hadoop? 18. Что является основным ядром всех продуктов Hadoop? 19. Что такое контрольная точка? 20. Как работает Hadoop MapReduce? 21. Что происходит при сбое узла данных? 22. Объясните, что такое спекулятивное исполнение 23. Укажите, в чем разница между СУБД и Hadoop

№ КМ	Вид КМ	Наименование КМ	Оценочные средства
			24. Объясните, как можно отлаживать код Hadoop 25. Объясните, что такое узлы хранения и вычисления 26. Укажите принципиальные различия между реляционной базой данных и HDFS?
8.	Текущий контроль	ЛР 4 Spark	<p>Цель: Ознакомиться фреймворком Apache Spark для реализации распределённой обработки неструктурированных и слабо структурированных данных, входящий в экосистему проектов Hadoop.</p> <p>Задание Напишите программу на Python, используя фреймворк Apache Spark, согласно вашему варианту для датасета "brooklyn_sales_map.csv". Для помощи в выполнении задания можно воспользоваться книгами «Изучаем Spark. Молниеносный анализ данных -2015», «PySpark SQL Recipes» или интернетом.</p> <p>Пример индивидуального задания.</p> <ul style="list-style-type: none"> • Найдите среднюю площадь жилья (gross_sqft) и выведите новую таблицу, содержащую стоимость жилья и отклонение стоимости от среднего значения. • Найдите среднюю площадь жилья (gross_sqft) для каждого года, в котором оно было построено (year_built). • Найдите среднюю стоимость жилья (sale_price) для всех сочетаний соседств (neighborhood) и категорий класса здания (building_class_category). • В исходном датафрейме удалите все строки с записями домов, которые были построены позже 2000 года (year_built), а также те, которые содержат только нулевые значения. <p>Контрольные вопросы</p> <ol style="list-style-type: none"> 1. Что такое архитектура Apache Spark? 2. Что вы понимаете под трансформациями в Spark? 3. Как реализована потоковая передача в Spark? 4. Проиллюстрируйте некоторые недостатки использования Spark. 5. Объясните пять V больших данных 6. Как работает Spark Streaming? 7. Как реализована потоковая передача в Spark? 8. Какие существуют типы преобразования в DStream? 9. Hadoop использует репликацию для обеспечения отказоустойчивости. Как это достигается в Apache Spark? 10. Расскажите об основных функциях Apache Spark. 11. Что такое наборы данных Spark? 12. Что такое YARN в Spark? 13. Каковы преимущества Spark по сравнению с MapReduce? 14. Есть ли польза от изучения MapReduce, если Spark лучше, чем MapReduce? 15. Что такое перетасовка в Spark? Когда это происходит? 16. Что такое приемники в Apache Spark Streaming? 17. Предоставляет ли Apache Spark контрольные точки? 18. Что такое скалярные и агрегатные функции в Spark SQL? 19. Зачем нужны широковебательные переменные при работе с Apache Spark? 20. Перечислите несколько вариантов использования, в которых Spark превосходит Hadoop в обработке.

№ КМ	Вид КМ	Наименование КМ	Оценочные средства
			21. Какие языки поддерживает Apache Spark для разработки приложений, выполняющих обработку больших данных?
9.	Курсовая работа/проект	Защита курсовой работы	<p>Контрольные вопросы</p> <ol style="list-style-type: none"> 1. Поясните, почему база данных является моделью 2. Каковы состав и азначение уровней архитектуры ANSI/SPARC? 3. Какие основные функции должна выполнять СУБД? 4. Какин архитектурные решения досткпа к БД вам известны? 5. Почему архитектура файл-сервер не подходит для многопользовательских БД? 6. Как могут распределяться задачи между коентом и сервером БД? 7. Какие связи между сущностями могут быть отражены реляционной моделью? 8. Что понимается под термином «тип данных»? 9. Для чего предназначены доменные огреничения? 10. Для чего предназначен первичный ключ? 11. Какин требования предъявляются к реляционной таблице? 12. Каким образом организуется связь между двумя отношениями? 13. Что в реляционной модели понимается под целостностью данных? 14. Какие виды целостности данных вам известны? 15. Почему говорят, что реляционные БД работают в трехзначной логике? 16. Укажите место БД в циклн\е преобразования информации. 17. Какие этапы входят в жизненный циел БД? 18. В чем отличие между стратегиями восходящего и нисходящего проектирования? 19. Что понимается под концептуальным проектированием БД? 20. Какую роль в процессе проектирования БД играют высокоуровневые модели данных? 21. Какие типы связей поддерживает ER-модель? 22. Какие CASE-системы позволяют создавать ER-модели БД? 23. С какими анамалиями мы можем столкнуться при модификации данных в ненормализованной таблице? 24. На каком этапе проектирования БД осуществляется нормализация? 25. Почему процесс нармализации считается восходящим проектированием? 26. Дайте определение первой, второй, третьей нармальной форм и нормальной формы Бойса-Кодда? 27. Для чего необходимо индексирование таблиц? 28. Какие индексы в БД создаются автоматически? 29. Какие виды индексов вам известны? 30. Что понимается под угрозой информационной системе? 31. Что понимается под безопасностью данных?

№ КМ	Вид КМ	Наименование КМ	Оценочные средства
			<p>32. Раскройте смысл терминов: «конфиденциальность», «доступность», «целостность данных».</p> <p>33. Что на ваш взгляд является источником угроз для БД?</p> <p>34. Какие специфичные (присущие только БД) угрозы вам известны?</p> <p>35. Дайте определение политики безопасности.</p> <p>36. Что понимается под терминами: «идентификация», «аутентификация», авторизация»?</p> <p>37. Расскажите о правилах защиты БД.</p> <p>38. Что следует понимать под экономической оправданностью при организации защиты БД?</p> <p>39. Расшифруйте аббревиатуру DDL</p> <p>40. Какие задачи решаются с помощью доменов?</p> <p>41. Какие возможности по модификации таблиц предоставляет DDL?</p> <p>42. Каким образом при создании таблиц описываются правила обеспечения ссылочной целостности?</p> <p>43. В каком случае индексы таблиц создаются автоматически?</p> <p>44. Почему представление называют виртуальным отношением?</p> <p>45. Какие преимущества дают хранимые процедуры и функции?</p> <p>46. Почему вызов триггеров осуществляется автоматически?</p> <p>47. На какие события должен учесть реагировать триггер?</p> <p>48. Для чего предназначены курсоры?</p> <p>49. Что такое транзакция?</p> <p>50. Какие требования предъявляются к транзакциям?</p> <p>51. Для чего ведется журнал транзакций?</p> <p>52. В каких состояниях может находиться транзакция?</p> <p>53. Как работает метод блокировок?</p> <p>54. Как можно выйти из ситуации взаимной блокировки транзакций?</p> <p>55. Какие фазы содержит метод двухфазной блокировки?</p> <p>56. Какие требования предъявляются стандартом SQL к изоляции транзакций?</p> <p>57. Какие объекты БД нуждаются в защите?</p> <p>58. Дайте определение понятию «привилегия».</p> <p>59. Что такое «роль», и как она участвует в процедуре надления пользователя полномочиями?</p> <p>60. Какими возможностями обладает инструкция GRANT?</p> <p>61. Какими возможностями обладает инструкция REVOKE?</p>
10.	Промежуточная аттестация	Экзамен	<p>Вопросы к экзамену по курсу «Администрирование и проектирование хранилищ больших данных»</p> <ol style="list-style-type: none"> 1. Концепции баз, банков, хранилищ данных <ol style="list-style-type: none"> 1.1. Жизненный цикл БД 1.2. Постановка задачи разработки, инфологическая модель 1.3. Концептуальное проектирование и ER-модель 1.4. Логическое проектирование. Нормализация 1.5. Безопасность данных в PostgreSQL

№ КМ	Вид КМ	Наименование КМ	Оценочные средства
			<p>1.6. Работа с транзакциями в PostgreSQL. Блокировки. Уровни изоляции транзакций</p> <p>1.7. Использование индексов в PostgreSQL</p> <p>1.8. Оптимизация запросов для достижения максимальной производительности</p> <p>1.9. Резервное копирование и восстановление в PostgreSQL</p> <p>1.10. Резервное копирование и репликация в PostgreSQL</p> <p>2. Типы современных СУБД</p> <p>2.1. Типы современных СУБД</p> <p>2.2. CAP-теорема</p> <p>2.3. СУБД MongoDB. Особенности. Область применения.</p> <p>2.4. Клиент-серверные СУБД</p> <p>2.5. Модели распределения функций</p> <p>2.6. Распределенные БД. Правила Дейта. Фрагментация и репликация</p> <p>2.7. Распределенные транзакции</p> <p>2.8. Преимущества и недостатки распределенных СУБД</p> <p>2.9. Документо-ориентированные БД</p> <p>2.10. Подход NoSQL</p> <p>2.11. БД ключ-значение</p> <p>2.12. New SQL БД</p> <p>3. OLAP-технологии, хранилища больших данных</p> <p>3.1. Модель сервера баз данных. Достоинства и недостатки.</p> <p>3.2. Модель сервера приложений. Достоинства и недостатки</p> <p>3.3. Архитектура хранилищ данных</p> <p>3.4. OLAP-куб. Схемы звезда и снежинка.</p> <p>3.5. Требования к OLAP-инструментам</p> <p>4. Технологии работы с хранилищами больших данных</p> <p>4.1. Что такое «большие данные»</p> <p>4.2. Принципы работы с большими данными</p> <p>4.3. Архитектура Apache Hadoop</p> <p>4.4. Архитектура Spark</p> <p>4.5. Распределенная обработка MapReduce</p>

Пример экзаменационного билета

Билет № 3

1. Концептуальное проектирование и ER-модель
2. Преимущества и недостатки распределенных СУБД
3. Принципы работы с большими данными

Преподаватель _____
(подпись)

Заведующий кафедрой _____
(подпись)

