Министерство науки и высшего образования Российской Федерации

Федеральное государственное бюджетное образовательное учреждение высшего образования
«Сибирский государственный индустриальный университет»

Кафедра прикладных информационных технологий и программирования

УТВЕРЖДАЮ
Директор института
информационных технологий и
автоматизированных систем
_____ А.В. Зимин
подпись
«____» _____ 20___ г.

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ

Постреляционные базы данных

09.03.03 «Прикладная информатика»
(направленность (профиль): «Искусственный интеллект и машинное обучение»)

Квалификация выпускника
Бакалавр

Форма обучения
Очная форма

Срок обучения: 4 года

Год начала подготовки 2026

Новокузнецк
2026

СибГИУ
Подлинник электронного документа хранится в системе электронного согласования документов СибГИУ
ДОКУМЕНТ ПОДПИСАН ЭП
Сертификат: daba56d7b892efca3bf5c04ed0251c98b19b099c
Владелец: Директор ИИТиАС Зимин Алексей Валерьевич
Действителен: с 15 января 2025 г. по 10 апреля 2026 г.
Дата и время: 10-03-2026 19:28:10 ПРОВЕРИТЬ ПОДПИСЬ

1 Цели и задачи освоения учебной дисциплины

Целями учебной дисциплины являются:

– формирование системного представления, первичных знаний, умений и навыков обучающихся по проектированию и использованию постреляционных баз данных.

Задачами учебной дисциплины являются:

– формирование представлений о роли и месте постреляционных баз данных в информационных системах, о назначении и основных характеристиках различных типов постреляционных баз данных, их функциональных возможностях;

– формирование практических навыков и умений по созданию и сопровождению постреляционных баз данных.

2 Место учебной дисциплины в структуре ООП по направлению подготовки (специальности)

Учебная дисциплина относится к учебным дисциплинам части, формируемой участниками образовательных отношений **Блока 1 «Дисциплины (модули)»** ООП по направлению подготовки (специальности) 09.03.03 «Прикладная информатика».

Учебная дисциплина базируется на предварительном усвоении обучающимися учебных дисциплин:

– Базы данных;

– Основы проектной деятельности.

Учебная дисциплина дополняет знания и умения, получаемые по одновременно изучаемым и последующим учебным дисциплинам:

– Проектирование информационных систем;

– Разработка API сайтов и работа с web-фреймворками;

– Проектная деятельность 2;

– Выполнение и защита выпускной квалификационной работы.

3 Планируемые результаты обучения по учебной дисциплине

Процесс изучения учебной дисциплины направлен на формирование следующих компетенций:

– Профессиональные компетенции

Наименование категории (группы) ПК	Код и наименование ПК	Код и наименование индикатора достижения ПК	Планируемые результаты обучения
	ПК-1: Способен участвовать в	ПК-1.1 Участвует в разработке	– знать: критерии выбора типа постреляционной базы данных в зависимости от характера данных

	разработке прототипов информационных систем	прототипа информационной системы в соответствии с требованиями	(структурированные/неструктурированные) и требований к масштабируемости. – уметь: анализировать требования к информационной системе и выбирать оптимальный тип постреляционной базы данных (MongoDB, Neo4j, Redis и др.) для реализации прототипа.
		ПК-1.2 Тестирует прототип информационной системы на корректность принятых решений	– знать: методы верификации данных и проверки запросов к NoSQL-базам данных. – уметь: выполнять тестовые запросы к разработанной постреляционной базе данных для проверки корректности выборки и модификации данных.

4 Объем и содержание учебной дисциплины

Учебные занятия по учебной дисциплине проводятся в форме контактной работы и в форме самостоятельной работы обучающихся.

Контактная работа включает в себя занятия лекционного типа (лекции), занятия семинарского типа (лабораторные работы, коллоквиумы), промежуточную аттестацию обучающихся и иные формы взаимодействия обучающихся с педагогическими работниками и (или) лицами, привлекаемыми к реализации ООП на иных условиях, в том числе при проведении промежуточной аттестации обучающихся. Контактная работа может проводиться с применением электронного обучения, дистанционных образовательных технологий.

Объем учебной дисциплины

Семестр / курс		***ИТОГО***	***5 семестр***
Форма промежуточной аттестации			*экзамен, зачет с оценкой по КР*
Трудоёмкость	*академ. час.*	***216***	*216*
	зачетных единиц	***6***	*6*
Лекции, *академ. час.*		***32***	*32*
в форме практической подготовки		***0***	*0*
Лабораторные работы, *академ. час.*		***32***	*32*
в форме практической подготовки		***0***	*0*
Практические занятия, *академ. час.*		***0***	*0*
в форме практической подготовки		***0***	*0*
Курсовая работа, *академ. час.*		***36***	*36*
в форме практической подготовки		***0***	*0*
Консультации, *академ. час.*		***0***	*0*
в форме практической подготовки		***0***	*0*
Самостоятельная работа, *академ. час.*		***62***	*62*

в форме практической подготовки	***0***	*0*
Контроль, *академ. час.*	***54***	*54*
в форме практической подготовки	***0***	*0*

Содержание учебной дисциплины

Раздел 1 Введение в постреляционные БД;

Тема 1.1 Факторы появления постреляционных БД (Большие объемы данных, гибкость схемы, поддержка полуструктурированных данных и распределенные вычисления);

Тема 1.2 Классификация постреляционных баз данных (Документоориентированные, графовые и колоночные системы, системы ключ-значение);

Раздел 2 Системы управления БД NoSQL типа;

Тема 2.1 Системы ключ-значение (Хранение пар ключ-значение. Масштабируемость обработки и производительность для доступа к данным);

Тема 2.2 Колоночные системы данных (Колоночная структура данных, сжатие данных, скорость выполнения запросов);

Тема 2.3 Документоориентированные системы (Данные в формате BSON. Синтаксис языка запросов к документоориентированной БД);

Тема 2.4 Графовые системы (Данные в виде узлов и ребер, моделирование связей между данными);

Раздел 3 Распределенные системы обработки больших данных;

Тема 3.1 Большие данные (Характеристики Big Data (VVV). Источники и форматы данных. Проблемы хранения и обработки сверхбольших объемов информации);

Тема 3.2 Распределенные файловые системы, Hadoop (Архитектура HDFS (NameNode/DataNode). Отказоустойчивость и репликация данных. Экосистема и компоненты Hadoop);

Тема 3.3 Технология MapReduce (Модель параллельных вычислений. Стадии обработки (Map, Shuffle, Reduce).
Примеры реализации задач (WordCount, Join)).

5 Перечень тем лекций

№ раздела / темы дисциплины	Темы лекций	Трудоемкость, *академ.час*	
		всего	в форме практической подготовки
Раздел 1.	Введение в постреляционные БД	10	
Раздел 2.	Системы управления БД NoSQL типа	16	
Раздел 3.	Распределенные системы обработки больших данных	6	

Итого:	**32**	**0**

6 Перечень тем практических занятий (семинаров)

№ раздела / темы дисциплины	Темы практических занятий (семинаров)	Трудоемкость, *академ.час*	
		всего	**в форме практической подготовки**
	Отсутствуют		
Итого:		**0**	**0**

7 Перечень тем лабораторных работ

№ раздела / темы дисциплины	Темы лабораторных работ	Трудоемкость, *академ.час*	
		всего	**в форме практической подготовки**
Раздел 1.	Анализ требований и выбор оптимальной NoSQL-системы для решения прикладных задач	4	
Тема 2.1.	Работа с хранилищами ключ-значение	7	
Тема 2.2.	Работа с колоночными БД	7	
Тема 2.3.	Работа с документоориентированными БД	7	
Тема 2.4.	Работа с графовыми БД	7	
Итого:		**32**	**0**

8 Перечень тем курсовых работ (проектов)

№ раздела / темы дисциплины	Темы курсовых работ (проектов)	Трудоемкость, *академ.час*	
		всего	**в форме практической подготовки**
Раздел 1; Раздел 2; Раздел 3.	Проектирование и разработка базы данных для хранения постов и комментариев в блоге. Проектирование и разработка базы данных каталога книг с фильтрацией по жанру, автору и году. Проектирование и разработка базы данных для органайзера (заметки, задачи, теги). Проектирование и разработка базы данных рецептов с поиском по	36	

	ингредиентам. Проектирование и разработка базы данных для счетчика посещаемости веб-страниц. Проектирование и разработка базы данных для управления сессиями пользователей. Проектирование и разработка базы данных списка последних просмотренных товаров		
Итого:		**36**	**0**

9 Виды самостоятельной работы

№ раздела / темы дисциплины	**Виды самостоятельной работы**	**Трудоемкость, *академ.час***	
		всего	**в форме практической подготовки**
Раздел 1.	1. Изучение лекционного материала; 2. Подготовка к лабораторной работе; 3. Прохождение тестирования.	21	
Раздел 2.	1. Изучение лекционного материала; 2. Подготовка к лабораторной работе; 3. Прохождение тестирования.	21	
Раздел 3.	1. Изучение лекционного материала; 2. Прохождение тестирования.	20	
Курсовая работа	*Выполнение курсовой работы*	36	0
Контроль	*Подготовка к экзамену*	54	
Итого:		**152**	**0**

10 Матрица компетентностных задач по учебной дисциплине

Контролируемые разделы (темы) дисциплины	**Контролируемые компетенции**	**Оценочные средства**
Раздел 1.	ПК-1	1. Наблюдение за работой на лабораторной работе; 2. Оценивание этапов выполнения курсовой работы; 3. Тестовые задания

Раздел 2.	ПК-1	1. Наблюдение за работой на лабораторной работе; 2. Оценивание этапов выполнения курсовой работы; 3. Тестовые задания
Раздел 3.	ПК-1	1. Оценивание этапов выполнения курсовой работы; 2. Тестовые задания
Курсовая работа.	ПК-1	1. Вопросы к защите курсовой работы
Экзамен.	ПК-1	1. Вопросы к экзамену

11 Входной контроль

11.1 Цель входного контроля − определить начальный уровень подготовленности обучающихся и выстроить индивидуальную траекторию обучения. В условиях личностно-ориентированной образовательной среды результаты входного оценивания обучающегося используются как начальные значения в индивидуальном профиле академической успешности обучающегося.

11.2 Описание оценочных средств

Форма проведения входного контроля – тестирование.

Длительность тестирования – 30 минут.

Тестирование проводится в течение первых двух недель изучения учебной дисциплины в СУО СибГИУ.

Для проведения тестирования используются тестовые задания для входного контроля (для ответа на 1 вопрос, как правило, выделяется 1 минута).

Шкала оценивания результатов тестирования

За каждый правильный ответ выставляется 1 балл. Максимальное количество – 10 баллов.

Оценка формируется в соответствии с критериями таблицы:

Максимальный балл	Проходной балл	Оценка
10	не менее 9	отлично
8	не менее 7	хорошо
6	не менее 5	удовлетворительно
4	-	неудовлетворительно

Основываясь на результатах тестирования, формируется индивидуальная траектория обучающегося (группы обучающихся), исходя из следующих соображений:

– 7 – 10 баллов – обучающийся готов к изучению учебной дисциплины без дополнительного повторения изученного ранее материала;
– 5 – 6 баллов – обучающийся в целом готов к изучению учебной дисциплины, требуется повторение отдельных вопросов(тем) изученного ранее материала, необходимого для успешного освоения учебной дисциплины;
– менее 5 баллов – обучающийся имеет слабую подготовку к изучению учебной дисциплины, необходимы корректирующие занятия и консультации.

Тестовые задания для входного контроля знаний обучающихся приведены в приложении.

12 Текущий контроль

12.1 Текущий контроль знаний используется для оперативного и регулярного управления учебной деятельностью (в том числе самостоятельной) обучающихся. Данный вид контроля стимулирует у обучающихся стремление к систематической самостоятельной работе по изучению учебной дисциплины и позволяет оценить степень освоения изученного материала.

Текущий контроль успеваемости осуществляется в течение семестра, в ходе повседневной учебной работы по индивидуальной инициативе педагогического работника, в том числе посредством использования СУО СибГИУ.

Формы проведения текущего контроля – наблюдение за работой на лабораторной работе; оценивание этапов выполнения курсовой работы; тестовые задания.

12.2 Описание оценочных средств

Шкалы оценивания (методики оценки)

12.2.1 Рекомендации по оцениванию этапов выполнения курсовой работы

Курсовая работа – конечный продукт, получаемый в результате планирования и выполнения комплекса учебных и исследовательских заданий. Позволяет оценить умения обучающихся самостоятельно конструировать свои знания в процессе решения практических задач и проблем, ориентироваться в информационном пространстве, а также уровень сформированности аналитических, исследовательских умений, умений практического и творческого мышления.

Оценивание этапов выполнения курсовой работы проводится в рамках текущей аттестации (три раза в семестр) в соответствии с графиком выполнения курсовой работы по учебной дисциплине. Результаты аттестации обучающихся по выполнению ими курсовой

работы оцениваются руководителями в процентах от общего объема выполненной работы. Сроки выполнения курсовой работы устанавливаются кафедрой в графике выполнения курсовой работы. График выполнения курсовой работы выдается старостам групп в начале каждого семестра.

Критерии оценки % выполнения курсовой работы

Этап работы	% выполнения	Форма отчетности
Для выбранной предметной области необходимо выполнить анализ требований и выбор оптимальной NoSQL-системы	30	Файл с выполненным заданием загружается в СУО СибГИУ
Для выбранной предметной области реализовать спроектированную базу данных средствами выбранной постреляционной СУБД	60	Файл с выполненным заданием загружается в СУО СибГИУ
Для выбранной предметной области предусмотреть возможность поиска данных, оптимизировать работу базы данных	100	Файл с выполненным заданием загружается в СУО СибГИУ

Критерии оценки курсовой работы

Критерии оценки выполнения курсовой работы	Оценка
Задание выполнено полностью: цель достигнута; основные понятия выделены; имеются в наличии схемы, графическое выделение особо значимой информации; все выводы и рекомендации обоснованы, работа тщательно вычитана, отсутствуют грамматические и стилистические ошибки, работа выполнена в полном объёме; процент выполнения задания соответствует плановому.	отлично
Задание выполнено: цель выполнения достигнута; наличие правильных эталонных ответов; однако работа выполнена не в полном объёме, имеются единичные опечатки, орфографические, пунктуационные, грамматические, стилистические ошибки; график выполнения задания нарушен на 10 – 20 %.	хорошо
Задание выполнено частично: цель выполнения достигнута не полностью; многочисленные ошибки снижают качество выполненной работы; работа представлена с отклонением от запланированного срока на 50 %.	удовлетворительно
Задание не выполнено, цель работы не достигнута.	неудовлетворительно

12.2.2 Рекомендации по оцениванию результатов тестирования обучающихся

Тестирование проводится по изученным разделам (темам) учебной дисциплины для выявления уровня знаний и умений обучающегося, мотивирования обучающегося к активизации самостоятельной работы по усвоению учебного материала, выявления и устранения пробелов в знаниях, повышения дисциплины и организации учебной деятельности обучающегося.

Тестирование проводится в СУО СибГИУ в рамках часов, отведенных на самостоятельную работу обучающихся. Для тестирования используется 20 вопросов из базы типовых тестовых заданий по изученным разделам (темам) учебной дисциплины. Обучающийся выбирает один ответ из предложенных вариантов ответа. В случае выбора неправильного ответа обучающийся получает за данный вопрос 0 баллов; в случае правильного ответа – 1 балл.

Оценка формируется в соответствии с критериями таблицы:

Максимальный балл	Проходной балл	Оценка
100%	50%	Зачтено
49%	-	Не зачтено

12.2.3 Рекомендации по оцениванию в процессе наблюдения за работой на лабораторном занятии

Наблюдение за работой на лабораторном занятии – важнейший компонент контроля деятельности обучающихся и качества их обучения.

Критерии оценки в процессе наблюдения за работой на лабораторном занятии:

– самостоятельность выполнения задания и работы с методическими указаниями (учитывается работа в течение занятия, быстрота и способность нахождения в методических указаниях нужной информации);

– правильность выполнения работы (учитывается последовательность выполняемых действий, умение работать с инструментами и/или средствами вычислительной техники);

– рациональность использования времени, отведенного на задание (не одобряется затянутость выполнения работы с учетом индивидуальных особенностей обучающихся);

– активность обучающегося при выполнении лабораторной работы (учитывается индивидуальная работа и работа в группе).

Критерии оценки по результатам наблюдения за работой на лабораторном занятии

Наименование параметра оценки	Максимальное количество баллов
Самостоятельно выполняет задания; умеет работать с методическими указаниями.	2,5
Указанные в методических указаниях действия выполняет в	1,25

Наименование параметра оценки	Максимальное количество баллов
правильной последовательности; умеет работать с инструментами и/или средствами вычислительной техники.	
Рационально использует время, отведенное на задание; проявляет активность при выполнении лабораторной работы.	1,25
Итого	**5**

Максимальное количество – 5 баллов соответствует оценке «отлично», 4 балла – «хорошо», 3 балла – «удовлетворительно», 2 балла – «неудовлетворительно».

Задания и вопросы для текущего контроля знаний обучающихся по разделам (темам) учебной дисциплины приведены в приложении.

13 Промежуточная аттестация

13.1 Оценочные материалы для промежуточной аттестации обучающихся по учебной дисциплине предназначены для оценки степени достижения запланированных результатов обучения по завершению изучения учебной дисциплины в установленной учебным планом форме и позволяют определить качество усвоения изученного материала.

Итоговой формой контроля сформированности компетенций у обучающихся по учебной дисциплине является: курсовая работа; экзамен.

Оценочные материалы для промежуточной аттестации обучающихся включают: вопросы к защите курсовой работы; вопросы к экзамену.

13.2 Оценивание обучающегося на защите курсовой работы

Оценка	Требования к знаниям
отлично	Оценка «отлично» выставляется обучающемуся, если выполнены все требования к выполнению и защите курсовой работы: соблюдены требования к внешнему оформлению, последовательно изложен материал; при защите обучающийся демонстрирует умения применять теоретические знания основной и дополнительной литературы, может показать и объяснить применение технологических решений, использованных в курсовой работе; даны правильные ответы на вопросы, заданные в процессе защиты курсовой работы.
хорошо	Оценка «хорошо» выставляется обучающемуся, если основные требования к выполнению и защите курсовой работы выполнены, но при этом допущены недочеты. В частности, имеются неточности в изложении материала; отсутствует логическая последовательность в решениях, имеются упущения в оформлении; на дополнительные вопросы при защите курсовой работы даны неполные ответы.

Оценка	Требования к знаниям
удовлетворительно	Оценка «удовлетворительно» выставляется обучающемуся, если имеются существенные отступления от требований к выполнению и защите курсовой работы. В частности: задача решена лишь частично; допущены фактические ошибки в содержании и оформлении; при ответе на дополнительные вопросы на защите курсовой работы обучающийся показывает знания только основного материала, испытывает затруднения при объяснении выбранных решений.
неудовлетворительно	Оценка «неудовлетворительно» выставляется обучающемуся, если задача решена лишь частично; допущены грубые ошибки в содержании и при оформлении или при ответе на дополнительные вопросы; во время защиты курсовой работы обучающийся обнаруживает существенное непонимание проблемы.

Если допущены существенные недостатки в оформлении курсовой работы, опущен или не написан какой-либо раздел, имеются отступления от задания на курсовую работу, то такая курсовая работа возвращается обучающемуся на доработку.

13.3 Оценивание обучающегося на экзамене

Оценка	Требования к знаниям
отлично	Оценка «отлично» выставляется обучающемуся, если он глубоко и прочно усвоил программный материал, исчерпывающе, последовательно, четко и логически его излагает, умеет тесно увязывать теорию с практикой, свободно справляется с задачами, вопросами и другими видами применения знаний, причем не затрудняется с ответом при видоизменении заданий, использует в ответе материал различной литературы, правильно обосновывает принятое нестандартное решение, обладает приемами выполнения практических задач по формированию компетенций.
хорошо	Оценка «хорошо» выставляется обучающемуся, если он твердо знает материал, грамотно и по существу излагает его, не допуская существенных неточностей в ответе на вопрос, правильно применяет теоретические положения при решении практических вопросов и задач, обладает необходимыми умениями и приемами их выполнения.
удовлетворительно	Оценка «удовлетворительно» выставляется обучающемуся, если он имеет знания только основного материала, но не усвоил его деталей, допускает неточности, недостаточно правильные формулировки, испытывает сложности при выполнении практических работ и затрудняется связать теорию вопроса с практикой.
неудовлетворительно	Оценка «неудовлетворительно» выставляется обучающемуся, который не знает значительной части программного материала, неуверенно отвечает, допускает серьезные ошибки, не имеет представлений по методике выполнения практической работы. Как правило, оценка «неудовлетворительно» ставится обучающимся, которые не могут продолжить обучение без дополнительных занятий по

Оценка	Требования к знаниям
	данной учебной дисциплине.

Задания и вопросы для промежуточной аттестации обучающихся приведены в приложении.

14 Учебно-методическое и информационное обеспечение учебной дисциплины

а) литература:

1 Стасышин, В. М. Базы данных: технологии доступа : учебное пособие для вузов / В. М. Стасышин, Т. Л. Стасышина. — 2-е изд., испр. и доп. — Москва : Издательство Юрайт, 2024. — 164 с. — ISBN 978-5-534-08687-4. – URL: https://urait.ru/bcode/538921 (дата обращения: 08.03.2026);

2 Гасанов, Э. Э. Интеллектуальные системы. Теория хранения и поиска информации : учебник для вузов / Э. Э. Гасанов, В. Б. Кудрявцев. — 2-е изд., испр. и доп. — Москва : Издательство Юрайт, 2024. — 271 с. — ISBN 978-5-534-08684-3. – URL: https://urait.ru/bcode/537938 (дата обращения: 08.03.2026);

3 Стружкин, Н. П. Базы данных: проектирование. Практикум : учебное пособие для вузов / Н. П. Стружкин, В. В. Годин. — Москва : Издательство Юрайт, 2024. — 291 с. — ISBN 978-5-534-00739-8. – URL: https://urait.ru/bcode/537149 (дата обращения: 08.03.2026);

4 Парфенов, Ю. П. Постреляционные хранилища данных : учебное пособие для вузов / Ю. П. Парфенов ; под научной редакцией Н. В. Папуловской. — Москва : Издательство Юрайт, 2024. — 121 с. — ISBN 978-5-534-09837-2. – URL: https://urait.ru/bcode/539330 (дата обращения: 08.03.2026);

5 Маркин, А. В. Системы графовых баз данных. Neo4j : учебное пособие для вузов / А. В. Маркин. — Москва : Издательство Юрайт, 2024. — 303 с. — ISBN 978-5-534-13996-9. – URL: https://urait.ru/bcode/544089 (дата обращения: 08.03.2026);

6 Стружкин, Н. П. Базы данных: проектирование : учебник для вузов / Н. П. Стружкин, В. В. Годин. — Москва : Издательство Юрайт, 2024. — 477 с. — ISBN 978-5-534-00229-4. – URL: https://urait.ru/bcode/536006 (дата обращения: 08.03.2026);

7 Редмонд, Э. Семь баз данных за семь недель. Введение в современные базы данных и идеологию NoSQL : монография / Э. Редмонд, Д. Р. Уилсон. – Москва : ДМК-пресс, 2013. – 384 с. – ISBN 978-5-94074-866-3. – URL: https://www.studentlibrary.ru/book/ISBN9785940748663.html (дата обращения: 08.03.2026).

б) ресурсы информационно-телекоммуникационной сети «Интернет»:

1 Электронный каталог : сайт / Научно-техническая библиотека СибГИУ. – Новокузнецк, [199 –]. – URL: http://libr.sibsiu.ru. – URL: https://libr.sibsiu.ru;

2 Электронная библиотека // Научно-техническая библиотека СибГИУ : сайт. – Новокузнецк, [200 –]. – URL: http://library.sibsiu.ru/LibrELibraryFullText.asp. – Режим доступа: для авторизир. пользователей. – URL: https://library.sibsiu.ru/LibrELibraryFullText.asp;

3 Консультант студента : электронно-библиотечная система / ООО «КОНСУЛЬТАНТ СТУДЕНТА». – Москва, [200 –]. – URL: http://www.studentlibrary.ru. – Режим доступа: для авторизир. пользователей;

4 ЛАНЬ : электронно-библиотечная система : [коллекция «Единая профессиональная база знаний для технических вузов - Издательство Лань" ЭБС»] / ООО «Издательство ЛАНЬ». – Санкт-Петербург, [200 –]. – URL: http://e.lanbook.com. – Режим доступа: для авторизир. пользователей;

5 Образовательная платформа ЮРАЙТ / ООО «Электронное издательство ЮРАЙТ». – Москва, [200 –]. – URL: https://urait.ru. – Режим доступа: для авторизир. пользователей;

6 Университетская библиотека онлайн : электронно-библиотечная система / ООО «Директ-Медиа». – Москва, [200 –]. – URL: https://biblioclub.ru. – Режим доступа: для авторизир. пользователей. – URL: http://www.biblioclub.ru;

7 НАУЧНАЯ ЭЛЕКТРОННАЯ БИБЛИОТЕКА eLIBRARY.RU : база данных / ООО «НЭБ». – Москва, [200 –]. – URL: http://elibrary.ru. – Режим доступа: по подписке;

8 Электронные периодические издания ИВИС : универсальная база данных / ООО «ИВИС». – Москва, [200 –]. – URL: http://eivis.ru. – Режим доступа: по подписке.

в) лицензионное и свободно распространяемое программное обеспечение:

– 7-Zip;
– ABBYY FineReader;
– Adobe Acrobat Reader;
– Astra Linux Special Edition;
– JetBrains PyCharm Community Edition;
– Kaspersky Endpoint Security;
– Microsoft Office;
– Microsoft Windows;
– OnlyOffice;
– PyCharm;
– Zoom;

– Р7-Офис.

г) базы данных и информационно-справочные системы:

1 ГАРАНТ : справочно-правовая система / ООО «Правовой центр «Гарант». – Кемерово, [200 –]. – Режим доступа: компьютерная сеть Сиб. гос. индустр. ун-та.;

2 КонсультантПлюс : справочно-правовая система / ООО «Информационный центр АНВИК». – Новокузнецк, [199 –]. – Режим доступа: компьютерная сеть библиотеки Сиб. гос. индустр. ун-та.;

3 Техэксперт : информационно-справочная система / ООО «Группа компаний «Кодекс». – Кемерово, [200 –]. – Режим доступа: компьютерная сеть Сиб. гос. индустр. ун-та.

15 Материально-техническое обеспечение учебной дисциплины

Материально-техническое обеспечение учебной дисциплины включает учебные аудитории, оснащенные оборудованием, компьютерной техникой, и техническими средствами обучения, в том числе:

- учебную аудиторию для проведения занятий лекционного типа, оборудованную учебной доской, экраном и мультимедийным проектором;
- учебную аудиторию для проведения занятий семинарского типа (лабораторных работ), оснащенную учебной доской, компьютерной техникой с возможностью подключения к сети Интернет и обеспечением доступа в электронную информационно-образовательную среду, научно-техническую библиотеку СибГИУ;
- учебную аудиторию для выполнения курсовых работ;
- учебную аудиторию (помещения) для групповых и индивидуальных консультаций, текущего контроля и промежуточной аттестации;
- помещения для самостоятельной работы, оснащенные компьютерной техникой с возможностью подключения к сети «Интернет» и обеспечением доступа в электронную информационно-образовательную среду, научно-техническую библиотеку СибГИУ.

Рабочая программа дисциплины составлена в соответствии с требованиями ФГОС ВО по направлению подготовки (специальности) 09.03.03 «Прикладная информатика».

Составитель(и):

доцент Маслова Елена Владимировна (кафедра прикладных информационных технологий и программирования).

Рабочая программа дисциплины рассмотрена и утверждена на заседании кафедры.

Согласована:

Системный аналитик группы
системной архитектуры департамента
информационных технологий
ООО "СИ СОФТ"  Е.А. Шапошникова

Приложение

Тестовые задания для входного контроля знаний обучающихся

Выберите один правильный ответ.

a) База данных - это?

a) набор данных, собранных на одной дискете;
b) данные, предназначенные для работы программы;
c) совокупность взаимосвязанных данных, организованных по определенным правилам, предусматривающим общие принципы описания, хранения и обработки данных;
d) данные, пересылаемые по коммуникационным сетям.

b) В записи файла реляционной базы данных может содержаться:

a) исключительно однородная информация (данные только одного типа);
b) только текстовая информация;
c) неоднородная информация (данные разных типов);
d) только логические величин;
e) исключительно числовая информация;

c) Информационная система - это?

1. система, в которой хранится информация о состоянии сети INTERNET
2. комплекс аппаратно-программных средств, предназначенный для хранения й поиска информации;
3. совокупность базы данных и всего комплекса аппаратно-программных средств для ее хранения, изменения и поиска информации, для взаимодействия с пользователем;
4. система, которая предоставляет пользователю требуемую им информацию.

d) В иерархической базе данных совокупность данных и связей между ними описывается:

5. таблицей;
6. сетевой схемой;
7. древовидной структурой;
8. совокупностью таблиц.

e) Наиболее точным аналогом реляционной базы данных может служить:

e) неупорядоченное множество данных;
f) вектор;

g) генеалогическое дерево;
h) двумерная таблица.

f) Какой из вариантов не является функцией СУБД?
i) реализация языков определения и манипулирования данными
j) обеспечение пользователя языковыми средствами манипулирования данными
k) поддержка моделей пользователя
l) защита и целостность данных
m) координация проектирования, реализации и ведения БД

g) Система управления базами данных представляет собой программный продукт, входящий в состав:
a) прикладного программного обеспечения.
b) операционной системы;
c) уникального программного обеспечения;
d) системного программного обеспечения;
e) систем программирования;

h) Что обязательно должно входить в СУБД?
a) процессор языка запросов
b) командный интерфейс
c) визуальная оболочка
d) система помощи

i) Как называется набор хранимых записей одного типа?
a) хранимый файл
b) представление базы данных
c) ничего из вышеперечисленного
d) логическая таблица базы данных
e) физическая таблица базы данных

j) Основные требования, предъявляемые к базе данных?
a) адаптивность и расширяемость
b) восстановление данных после сбоев
c) распределенная обработка данных
d) контроль за целостностью данных
e) все ответы

Задания и вопросы для текущего контроля знаний обучающихся

по разделам (темам) учебной дисциплины

1 Тестирование по результатам освоения лекционного материала и самостоятельной работы

Тестирование проводится для определения степени усвоения материалов лекций, а также выполнения самостоятельной работы в рамках часов, отведенных на самостоятельную работу обучающихся, в СУО СибГИУ.

Перечень вопросов для тестирования:

Раздел 1

1. Что из перечисленного НЕ является фактором появления постреляционных (NoSQL) баз данных?

А. Рост объемов данных, генерируемых веб-приложениями

Б. Необходимость жесткой схемы данных с заранее определенными типами полей

В. Потребность в горизонтальном масштабировании на кластеры серверов

Г. Необходимость обработки полуструктурированных данных (JSON, XML)

2. Какой из вариантов перевода аббревиатуры NoSQL считается наиболее точным на сегодняшний день?

А. Non SQL (не SQL)

Б. Not Only SQL (не только SQL)

В. New SQL (новый SQL)

Г. No Structured Query Language (без структурированного языка запросов)

3. Какая компания опубликовала описание технологии BigTable, ставшей прообразом колоночных баз данных?

А. Amazon

Б. Facebook

В. Google

Г. Yahoo!

4. Какая модель данных использует хранение информации в виде пар «ключ-значение»?

А. Документоориентированная

Б. Графовая

В. Система «ключ-значение»

Г. Колоночная

5. Какой тип постреляционных баз данных лучше всего подходит для хранения и анализа социальных графов (связей между пользователями)?

А. Документоориентированные системы

Б. Графовые системы

В. Системы «ключ-значение»

Г. Колоночные системы

6. Какой формат данных чаще всего используется в документоориентированных базах данных?

А. CSV

Б. XML

В. JSON

Г. HTML

7. Redis является классическим представителем какого типа постреляционных баз данных?

А. Документоориентированные

Б. Графовые

В. Колоночные

Г. Системы «ключ-значение»

8. Какая характеристика НЕ относится к полуструктурированным данным?

А. Наличие тегов или маркеров для описания элементов

Б. Иерархическая или древовидная структура

В. Строгая табличная структура с фиксированным набором полей

Г. Возможность изменения структуры от записи к записи

9. MongoDB относится к какому типу постреляционных баз данных?

А. Системы «ключ-значение»

Б. Документоориентированные системы

В. Колоночные системы

Г. Графовые системы

10. Apache Cassandra и HBase являются примерами какого типа постреляционных баз данных?

А. Документоориентированные

Б. Графовые

В. Системы «ключ-значение»

Г. Колоночные системы

11. Какой тип масштабирования характерен для постреляционных баз данных в отличие от традиционных реляционных СУБД?

А. Вертикальное масштабирование (увеличение мощности одного сервера)

Б. Горизонтальное масштабирование (добавление новых серверов в кластер)

В. Масштабирование только за счет увеличения оперативной памяти

Г. Отсутствие возможности масштабирования

12. Neo4j является примером какого типа постреляционных баз данных?

А. Документоориентированные

Б. Графовые

В. Колоночные

Г. Системы «ключ-значение»

13. Какие типы данных хорошо поддерживаются постреляционными базами данных?

А. Строго структурированные табличные данные

Б. Полуструктурированные данные (JSON, XML)

В. Данные с заранее определенной схемой

Г. Неструктурированные данные (тексты, медиафайлы)

Д. Данные, требующие сложных JOIN-операций

14. Как расшифровывается аббревиатура NoSQL в современном понимании? (введите правильный ответ латиницей)

15. Какая теорема описывает компромиссы между согласованностью, доступностью и устойчивостью к разделению в распределенных системах? (введите аббревиатуру)

16. Установите соответствие между типами постреляционных баз данных и их характерными особенностями:

Тип базы данных	Особенность
1. Документоориентированные	А. Хранение данных в виде узлов и связей между ними
2. Графовые	Б. Хранение данных в виде коллекций документов (JSON)
3. Системы «ключ-значение»	В. Хранение данных в разреженных матрицах, оптимизация для аналитики
4. Колоночные	Г. Простейшая модель с доступом к значению по уникальному ключу

17. Какой формат данных, основанный на JavaScript, стал стандартом де-факто для документоориентированных баз данных? (введите аббревиатуру)

18. Какой тип постреляционных баз данных хранит данные в виде коллекций документов, каждый из которых может иметь свою структуру? (введите название)

19. Установите соответствие между названиями систем и их типами:

Название системы	Тип БД
1. Cassandra	А. Система «ключ-значение»
2. Neo4j	Б. Документоориентированная
3. Redis	В. Колоночная
4. MongoDB	Г. Графовая

20. Установите соответствие между требованием к базе данных и наиболее подходящим типом постреляционной БД:

Требование	Тип БД
1. Анализ социальных связей и рекомендации друзей	А. Системы «ключ-значение»
2. Быстрое кэширование данных по уникальному идентификатору	Б. Колоночные системы
3. Хранение каталога товаров с различными характеристиками	В. Графовые системы
4. Агрегация и аналитика огромных массивов данных (OLAP)	Г. Документоориентированные системы

Раздел 2

1. Какое основное свойство характеризует систему хранения данных типа «ключ-значение»?

А. Возможность сложных запросов и агрегирования данных

Б. Простота реализации и высокая производительность операций чтения-записи

В. Поддержка транзакций ACID

Г. Хранение структурированных данных

2. Какой принцип лежит в основе хранилищ типа «ключ-значение»?

А. Каждое значение хранится отдельно от ключа

Б. Значения хранятся последовательно друг за другом

В. Каждый ключ однозначно определяет одно значение

Г. Ключ используется только для сортировки значений

3. Какие операции поддерживаются системами типа «ключ-значение»?

А. SELECT, INSERT, UPDATE, DELETE

Б. GET, SET, DEL

В. JOIN, UNION, INTERSECT

Г. CREATE TABLE, DROP TABLE

4. Чем отличаются колонко-ориентированные системы от традиционных реляционных баз данных?

А. Тем, что каждая запись должна содержать одинаковое количество полей

Б. Их способностью хранить произвольное число столбцов в одной строке

В. Невозможностью гибкого изменения схемы данных

Г. Очевидным отсутствием оптимизации для аналитики

5. Какую проблему решают колонко-ориентированные системы данных?

А. Проблему низкой скорости записи данных

Б. Проблему высокой стоимости серверного оборудования

В. Проблема анализа большого объема данных

Г. Вопрос о невозможности параллельной обработки данных

6. Какая ключевая характеристика отличает колонко-ориентированную базу данных от реляционной?

А. Упрощённая схема данных

Б. Горизонтальное разделение таблицы по ряду записей

В. Масштабируемость при обработке аналитических нагрузок

Г. Неприменение индексов

7. Основным отличием документо-ориентированных систем от реляционных является…

А. Наличие строгих схем данных

Б. Использование иерархической структуры документов

В. Требование фиксированной длины полей

Г. Применение сложных связей между таблицами

8. Какой формат данных наиболее часто применяется в документо-ориентированных системах?

А. XML

Б. CSV

В. Binary Blobs

Г. JSON

9. Чем характеризуется документо-ориентированное хранение данных?

А. Жесткими ограничениями на размер документа

Б. Возможность хранить объекты различной структуры вместе

В. Необходимостью жесткой привязки данных к схеме

Г. Трудностью адаптации к изменениям требований бизнеса

10. Чем отличается документо-ориентированная база данных от реляционной?

А. Отсутствие ограничений на изменение структуры документа

Б. Применением встроенных процедур

В. Невозможностью распределённого хранения данных

Г. Постоянством размера данных

11: Какое важное качество документо-ориентированных систем?

А. Очень низкая латентность

Б. Максимальная простота администрирования

В. Высокая устойчивость к сбоям

Г. Удобство моделирования данных сложной структуры

12. Что представляет собой графовая база данных?

А. Система, основанная на древовидной структуре данных

Б. Хранилище, организованное вокруг узлов и ребер
В. Реляционная база данных с поддержкой графов
Г. Логическое представление файлового формата JSON
13. К Основным компонентам графа относятся…
А. Узел (вершина)
Б. Ребро (связь)
В. Атрибут узла
Г. Все варианты верны
14. Преимущество графовых систем заключается в возможности…
А. Эффективно анализировать взаимосвязанные данные
Б. Быстро производить арифметические вычисления
В. Масштабироваться вертикально без потерь производительности
Г. Хранить небольшие объёмы данных
15. Чем полезны графовые базы данных?
А. Оптимальны для OLAP-анализа
Б. Идеально подходят для социального графического анализа и рекомендательных систем
В. Лучше работают с редкими видами данных
Г. Помогают ускорить обработку изображений
16. Могут ли разные документы внутри одной коллекции иметь различную структуру?
А. Нет, все документы обязаны соответствовать единой схеме
Б. Да, но это снижает производительность
В. Да, каждый документ автономен и может отличаться от остальных
Г. Такая возможность доступна лишь в определённых продуктах
17. Какое дополнительное свойство полезно в графовых базах данных?
А. Одновременная модификация разных частей графа
Б. Локализация проблем с производительностью
В. Простота миграции данных из реляционных систем
Г. Возможность учёта веса связи между узлами
18. Чем отличаются графовые базы данных от документо-ориентированных?
А. Грамотнее управляют повторяющимися данными
Б. Проще интегрируются с классическими реляционными системами
В. Предполагают чёткое определение модели данных
Г. Ориентированы на специфику анализа сетей и взаимосвязей
19. Какая СУБД служит ярким представителем графовых баз данных?
А. GraphQL
Б. RDF
В. Relational Database Management System (RDBMS)
Г. Neo4j

20. Что такое узел (vertex) в графовой базе данных?

А. Элемент данных, содержащий свойства и возможные связи с другими узлами

Б. Связь между двумя узлами

В. Уровень изоляции данных

Г. Метрика качества графа

Раздел 3

1. Какие из перечисленных характеристик традиционно относят к основным признакам Big Data (VVV)?

А. Volume, Veracity, Visualization

Б. Volume, Velocity, Variety

В. Variety, Value, Veracity

Г. Velocity, Variability, Visualization

2. Какие из перечисленных технологий/платформ являются ключевыми для хранения и обработки больших данных?

А. MySQL, Oracle Database

Б. Hadoop, Apache Spark

В. NoSQL базы данных (например, MongoDB, Cassandra)

Г. Microsoft Excel

3. Какие из следующих этапов входят в жизненный цикл анализа больших данных?

А. Сбор данных

Б. Предобработка данных

В. Визуализация данных

Г. Интерпретация результатов

4. Какова одна из главных проблем безопасности, связанных с Big Data?

А. Отсутствие данных для анализа

Б. Разрыв между объемом данных, требующих защиты, и объемом реально защищенных данных

В. Слишком низкая скорость обработки

Г. Отсутствие инструментов визуализации

5. Какой язык программирования НЕ используется как основной для работы с большими данными?

А. Python

Б. R

В. SQL

Г. Pascal

6. Что представляет собой Big Data в наиболее точном смысле?

А. Данные объемом более 1 ТБ

Б. Данные, которые невозможно обработать в Excel

В. Серия подходов, инструментов и методов обработки структурированных и неструктурированных данных огромных объемов

Г. Исключительно данные из социальных сетей

7. Каковы из самых популярных перечисленных сфер применения Big Data?

А. Маркетинг (персонализация, прогнозирование спроса)

Б. Логистика (оптимизация перевозок)

В. Медицина (анализ показателей, разработка лекарств)

Г. Госструктуры (статистика, безопасность)

8. Что является одной из ключевых задач этапа предобработки данных?

А. Очистка данных от ошибок и дубликатов

Б. Визуализация результатов

В. Внедрение решений в производство

Г. Нормализация данных

9. Какой из перечисленных методов анализа данных относится к машинному обучению?

А. Кластеризация

Б. Описательная статистика

В. Классификация

Г. Корреляционный анализ

10. Перечислите три основные характеристики Big Data, известные как «три V» (VVV), через запятую или пробел.

11. Какая модель параллельных вычислений лежит в основе работы платформы Hadoop? (введите название)

12. Какой тип аналитики направлен на выявление текущих состояний процессов и отвечает на вопрос «Что произошло?» (введите название на русском или английском)

13. Какая технология из экосистемы Hadoop предназначена для высокоскоростной обработки больших данных в памяти и поддерживает как пакетную, так и потоковую обработку? (введите название)

14. Закончите определение: Big Data — это серия подходов, инструментов и методов обработки структурированных и _____ данных огромных объемов. (введите словосочетание)

15. Установите соответствие между технологией и её ролью в экосистеме больших данных:

Технология	Роль
1. Hadoop	А. Распределенная файловая система и модель MapReduce для пакетной обработки
2. Apache Spark	Б. Альтернатива классическим СУБД для работы с неструктурированными данными
3. NoSQL базы данных	В. Фреймворк для высокоскоростной обработки данных в памяти

16. Установите соответствие между этапом жизненного цикла Big Data и его ключевой задачей:

Этап	Задача
1. Сбор данных	А. Очистка от ошибок, дубликатов и нормализация
2. Предобработка данных	Б. Получение информации из различных источников (сенсоры, соцсети, логи)
3. Анализ данных	В. Представление результатов в понятном графическом виде
4. Визуализация данных	Г. Применение методов машинного обучения и статистики

17. Установите соответствие между проблемой безопасности Big Data и её описанием:

Проблема	Описание
1. Разрыв защиты	А. Несоответствие между объемом данных, требующих защиты, и реально защищенным объемом
2. Обезличивание данных	Б. Принцип хранения персональных данных, удаляющий прямые идентификаторы личности
3. Шифрование	В. Преобразование данных в форму, недоступную для чтения без специального ключа

18. Установите соответствие между категорией или источником данных и примером:

Категория/источник	Пример
1. Неструктурированные данные	А. Видеофайлы с камер наблюдения
2. Полуструктурированные данные	Б. Данные о посещениях веб-сайтов в формате логов
3. Данные социальных сетей	В. JSON-файлы с информацией о пользователях

Категория/источник	Пример
4. Маркетинговые данные	Г. История покупок и предпочтения клиентов

19. Установите соответствие между требованием к технологиям Big Data и его характеристикой:

Требование	Характеристика
1. Высокая производительность	А. Способность обрабатывать запросы с минимальной задержкой
2. Масштабируемость	Б. Гарантия сохранности данных при сбоях оборудования
3. Надежность	В. Возможность наращивать мощность путем добавления новых узлов
4. Конфиденциальность	Г. Защита персональных данных от несанкционированного доступа

20. Установите соответствие между методом, инструментом или языком программирования и его ролью в обработке больших данных:

Метод/инструмент/язык	Роль
1. Python	А. Очистка данных от ошибок, дубликатов и пропусков
2. Кластеризация	Б. Язык программирования для статистического анализа и машинного обучения
3. Очистка данных	В. Метод машинного обучения для группировки объектов без заранее известных меток
4. Java	Г. Язык программирования, часто используемый для реализации горизонтально масштабируемых систем (Hadoop, Spark)

2 Выполнение лабораторных работ

Задания для лабораторных работ, теоретический и методический материал для самостоятельного изучения и вопросы для собеседования представлены в методических указаниях к лабораторным работам.

3 Вопросы к экзамену для промежуточной аттестации

1 Перечислите и охарактеризуйте основные факторы, обусловившие появление постреляционных (NoSQL) баз данных.
2 Раскройте понятие «гибкость схемы» (schema-less) и объясните, почему оно стало важным в современных условиях.
3 Какие ограничения реляционных баз данных проявляются при работе с большими объемами данных и распределенными системами?
4 Приведите классификацию NoSQL-баз данных по модели данных. Кратко охарактеризуйте каждый тип.
5 Сравните документоориентированные и графовые базы данных с точки зрения области применения.
6 В чем принципиальное отличие колоночных систем от систем «ключ-значение»?
7 Опишите базовую модель хранения данных в системах типа «ключ-значение». Приведите примеры таких систем.
8 За счет каких механизмов достигается высокая производительность и масштабируемость в системах «ключ-значение»?
9 В каких сценариях (задачах) наиболее эффективно использование хранилищ типа «ключ-значение»?
10 Объясните, что означает «колоночное хранение данных». Чем оно отличается от строчного (row-oriented) хранения?
11 Каким образом колоночное сжатие данных повышает эффективность выполнения аналитических запросов?
12 Для каких типов нагрузок (OLTP или OLAP) оптимизированы колоночные СУБД и почему?
13 Что представляет собой формат BSON? Каковы его преимущества по сравнению с обычным JSON при хранении в базе данных?
14 Опишите структуру документа в документоориентированной БД. Как выполняются запросы к таким данным (основные операторы)?
15 В чем разница между внедренными (embedded) и ссылочными (referenced) документами при моделировании связей?
16 Из каких основных элементов (концепций) состоит графовая модель данных?
17 Как в графовых базах данных моделируются сложные связи между сущностями? Приведите пример.
18 Для решения каких классов задач графовые базы данных превосходят реляционные по производительности запросов?
19 Раскройте содержание характеристик Big Data (VVV или 5V). Что означают Volume, Velocity, Variety?
20 Перечислите основные источники больших данных и форматы, в которых они могут быть представлены.

21 С какими ключевыми проблемами сталкиваются при хранении и обработке сверхбольших объемов информации?
22 Опишите архитектуру HDFS. Как распределены роли между NameNode и DataNode?
23 Объясните механизм обеспечения отказоустойчивости в HDFS за счет репликации данных.
24 Перечислите основные компоненты экосистемы Hadoop (не менее 3-4) и укажите их назначение.
25 Опишите модель параллельных вычислений MapReduce. Раскройте содержание и назначение каждой стадии обработки: Map, Shuffle, Reduce.

4 Выполнение курсовой работы

В рамках учебной дисциплины предусмотрена самостоятельная работа обучающихся, результатом выполнения которой является оформленная в соответствии с требованиями пояснительная записка курсовой работы по заданной теме.

Примерный перечень тем курсовых работ представлен в рабочей программе дисциплины. Задание и методические указания к выполнению курсовой работы размещены в СУО СибГИУ.

Перед выдачей темы курсовой работы обучающийся знакомится с графиком его выполнения и критериями оценки. Выполненная пояснительная записка курсовой работы загружается обучающимся в СУО СибГИУ.

5 Вопросы к защите курсовой работы

1 Опишите выбранную предметную область. В чем заключается актуальность вашей работы для этой предметной области?

2 Перечислите предпосылки к разработке базы данных, решающего поставленную задачу.

3 Какими методами производится проектирование базы данных?

4 Какие задачи решает разработанная база данных? Каков ее функционал?

5 Перечислите требования, предъявляемые к базе данных.

6 Какой вид постреляционной базы данных был выбран? Почему?

7 Как в вашей базе данных соблюдаются условие непротиворечивости и целостности?

8 Каким образом реализован доступ к данным в вашей базе данных?

9 Как осуществляется администрирование вашей БД. Обоснуйте.

10 Какие группы пользователей существуют в вашей БД? Чем обусловлен выбор?