

Министерство науки и высшего образования Российской Федерации
НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
ТОМСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ (НИ ТГУ)

Институт прикладной математики и компьютерных наук

УТВЕРЖДАЮ:
Директор



А. В. Замятин

« 15 » июня 20 23 г.

Рабочая программа дисциплины

Постреляционные модели данных и промышленные СУБД - I

по направлению подготовки

01.04.02 Прикладная математика и информатика

Направленность (профиль) подготовки :
Big Data and Data Science

Форма обучения
Очная

Квалификация
Магистр

Год приема
2023

Код дисциплины в учебном плане: Б1.О.04

СОГЛАСОВАНО:

Руководитель ОП

А.В. Замятин

Председатель УМК

С.П. Сущенко

Томск – 2023

1. Цель и планируемые результаты освоения дисциплины

Целью освоения дисциплины является формирование следующих компетенций:

- ОПК-4 – способность комбинировать и адаптировать существующие информационно-коммуникационные технологии для решения задач в области профессиональной деятельности с учетом требований информационной безопасности;
- ПК-5 – способность выбирать методы, оформлять техническое задание и разрабатывать алгоритмы решения задач анализа промышленных данных.

Результатами освоения дисциплины являются следующие индикаторы достижения компетенций:

ИОПК-4.2 Учитывает основные требования информационной безопасности.

ИПК-5.1 Использует современные технологии обработки информации, вычислительную технику при решении задач анализа промышленных данных.

ИПК-5.2 Умеет производить сбор промышленных данных, знает специфику таких данных.

ИПК-5.3. Оформляет техническое задание для задачи профессиональной области.

2. Задачи освоения дисциплины

- Знать современные инструменты долговременного хранения данных;
- Выбирать инструменты долговременного хранения данных для их практического применения;
- Владеть навыками использования теоретических и практических знаний в области построения реляционных баз данных;
- Проектировать модели реляционных и постреляционных баз данных.

3. Место дисциплины в структуре образовательной программы

Дисциплина относится к обязательной части образовательной программы.

4. Семестр(ы) освоения и форма(ы) промежуточной аттестации по дисциплине

Первый семестр, зачет.

5. Входные требования для освоения дисциплины

Для успешного освоения дисциплины требуются результаты обучения по следующим дисциплинам: «Математические методы и модели для компьютерных наук», «Алгоритмы и структуры данных», знание Python .

6. Язык реализации

Английский.

7. Объем дисциплины

Общая трудоемкость дисциплины составляет 3 з.е., 108 часов, из которых:

-лекции: 12 ч.

-лабораторные: 24 ч.

в том числе практическая подготовка: 0 ч.

Объем самостоятельной работы студента определен учебным планом.

8. Содержание дисциплины, структурированное по темам

Тема 1. История развития и становления концепции баз данных. Реляционные базы данных.

Историческое развитие систем хранения данных. Основные понятия баз данных, структур данных и систем управления базами данных. Реляционная модель. Основные понятия и термины реляционной модели. Реляционные СУБД. Понятие нормальной формы. SQL - стандартный язык запросов к реляционным СУБД. Понятие целостности данных. Обработка транзакций. Недостатки реляционной модели данных.

Тема 2. Объектно-ориентированные и постреляционные СУБД. Постреляционные базы данных(NoSQL). Темпоральные базы данных.

Объектно-ориентированные СУБД. Постреляционные базы данных(NoSQL). Теорема CAP. Основные виды баз данных NoSQL. Преимущества и недостатки баз данных NoSQL. Область применения. Основные принципы, лежащие в основе темпоральных моделей данных. Понятие времени в темпоральных моделях данных. Модели, используемые в темпоральных базах данных (TRM, HDM). Темпоральность в реляционной СУБД.

Тема 3. Документно-ориентированные базы данных.

Документно-ориентированная модель данных. Достоинства и недостатки модели, области применения. Методы работы с данными. Структуры хранения данных. База данных MongoDB. Поисковая система Elasticsearch.

Тема 4. Графовые базы данных

Теория графов. Графовая модель данных. Достоинства и недостатки модели, области применения. Методы работы с данными. Структуры хранения данных. База данных Neo4J. Запросы на языке Cypher.

Тема 5. Базы данных «ключ-значение».

Модель данных «ключ-значение». Достоинства и недостатки модели, области применения. Методы работы с данными. Структуры хранения данных. База данных Riak, Redis.

Тема 6. Колоночные базы данных.

Колоночные БД. Модель данных. Достоинства и недостатки модели, области применения. Методы работы с данными. Структуры хранения данных. База данных Cassandra.

9. Текущий контроль по дисциплине

Текущий контроль по дисциплине проводится путем контроля посещаемости, тестов по лекционному материалу, выполнения лабораторных работ, практических работ и фиксируется в форме контрольной точки не менее одного раза в семестр.

10. Порядок проведения и критерии оценивания промежуточной аттестации

Зачет в первом семестре проводится в письменной форме в виде теста, который содержит не менее 30 теоретических и практических вопросов. Продолжительность зачета 1,5 часа. Необходимым условием для допуска к зачетному тесту является успешное выполнение всех лабораторных и практических работ по курсу.

Пример тестового вопроса:

1. К какому типу баз данных noSQL относится MongoDB?

Ответы:

- a) БД «ключ-значение»
- b) Графовые БД
- c) Документно-ориентированные БД
- d) Хранилища семейств колонок.

Результаты зачетного теста определяются в баллах и составляют 30% рейтинга.

11. Учебно-методическое обеспечение

а) Оценочные материалы текущего контроля представляют собой наборы тестов по темам теоретического блока.

б) Методические указания по проведению лабораторных работ представлены в виде описания лабораторных работ в файлах pdf. Список лабораторных работ:

Лаб.1. Логическое и физическое проектирование в Oracle DataModeler

Лаб.2. Проектирование темпоральной модели данных.

Лаб.3. Работа с документно-ориентированной БД MongoDB. Создание запросов к данным.

Лаб.4. Создание графовой модели данных. Создание запросов к данным.

Лаб.5. Создание и работа с БД Redis (или Riak). Создание запросов к данным.

Лаб.6. Создание и работа с БД Cassandra. Создание запросов к данным.

12. Перечень учебной литературы и ресурсов сети Интернет

а) основная литература:

1.	Мартишин, С.А.	Базы данных. Практическое применение СУБД SQL- и NoSQL-типа для применения проектирования информационных систем Учебное пособие	2021
2.	Парфенов, Ю. П.	Постреляционные хранилища данных учебное пособие для вузов	2022
3.	Эрик, Р.	Семь баз данных за семь недель. Введение в современные базы данных и идеологию NoSQL	2013

б) дополнительная литература:

1.	Карпендер, Д.	Cassandra. Полное руководство	2017
2.	Брэдшоу, Ш.	Mongo DB Полное руководство	2020
3.	Нидхем, М.	Графовые алгоритмы	2020
4.	Робинсон, Я.	Графовые базы данных: новые возможности для работы со связанными данными	2016
5.	Карпендер, Д.	Cassandra. Полное руководство	2017
6.	Брэдшоу, Ш.	Mongo DB Полное руководство	2020

в) ресурсы сети Интернет:

Документация по работе с базами данных, представленная в открытом доступе на официальных сайтах.

1. <https://www.mongodb.com/>
2. <http://cassandra.apache.org/>
3. <https://neo4j.com/>
4. <https://www.elastic.co/>
5. <http://www.oracle.com/>

13. Перечень информационных технологий

а) лицензионное и свободно распространяемое программное обеспечение:

– Microsoft Office Standart 2013 Russian: пакет программ. Включает приложения: MS Office Word, MS Office PowerPoint,

– публично доступные облачные технологии Google Docs

– Открытые БД (Mongodb, Neo4j, Redis, Cassandra)

– Средство моделирования реляционных БД Oracle DataModeler

б) информационные справочные системы:

– Электронный каталог Научной библиотеки ТГУ –
<http://chamo.lib.tsu.ru/search/query?locale=ru&theme=system>
– Электронная библиотека (репозиторий) ТГУ –
<http://vital.lib.tsu.ru/vital/access/manager/Index>

14. Материально-техническое обеспечение

Для реализации дисциплины необходимы лекционные аудитории и аудитории для проведения лабораторных занятий с доступом в Internet. Специальные технические средства (проектор, компьютер и т.д.) требуются для демонстрации материала в рамках изучаемых разделов, проведения защиты проектов в конце семестра. Вся основная и дополнительная литература, необходимая для самостоятельной работы и подготовки к экзамену, имеется в научной библиотеке ТГУ. Для выполнения лабораторных работ необходимы базы данных развернутые на сервере института.

15. Информация о разработчиках

Мокина Елена Евгеньевна, старший преподаватель кафедры теоретических основ информатики.