

**Федеральное государственное автономное образовательное
учреждение высшего образования
«Московский физико-технический институт
(национальный исследовательский университет)»**

УТВЕРЖДЕНО

**Директор физтех-школы
прикладной математики и
информатики**

А.М. Райгородский

	Рабочая программа дисциплины (модуля)
по дисциплине:	Базы данных
по направлению:	Прикладная математика и информатика
профиль подготовки:	Искусственный интеллект и большие данные Сетевое обучение кафедра алгоритмов и технологий программирования
курс:	2
квалификация:	бакалавр

Семестр, формы промежуточной аттестации: 4 (весенний) - Экзамен

Аудиторных часов: 60 всего, в том числе:

лекции: 30 час.

семинары: 30 час.

лабораторные занятия: 0 час.

Самостоятельная работа: 54 час.

Подготовка к экзамену: 30 час.

Всего часов: 144, всего зач. ед.: 4

Программу составил: Н.А. Меркурьева, ассистент

Программа обсуждена на заседании кафедры алгоритмов и технологий программирования 30.08.2022

Аннотация

В курсе рассматриваются базы данных, история их появления и развития по сегодняшний день, а также особенности работы с ними. Прежде всего, изучаются предпосылки к созданию и теория, лежащая в основе появления первых реляционных баз, – реляционная алгебра. Происходит знакомство слушателей с понятиями реляционной алгебры, СУБД, языком SQL, а также формируется понимание отличий практической реализации от теории, лежащей в ее основе. Обозреваются наиболее распространенные в индустрии типы реляционных СУБД, изучаются их особенности, сходства и различия. Рассматриваются основные подходы и этапы проектирования баз данных. Вводится понятие нормальных форм и их разновидностей вплоть до НФБК. Изучаются сложные конструкции языка SQL: подзапросы, оконные функции, представления (view), хранимые процедуры и процедурное расширение языка SQL, триггеры. Рассматривается архитектура СУБД и жизненный цикл запроса на примере PostgreSQL. Обзорно рассматривается нереляционный подход в создании СУБД (NoSQL): их типы, конкретные примеры, особенности работы. Дается представление о современных применениях СУБД в индустрии.

Курс содержит в себе теоретическую базу, необходимую при работе с базами данных, в первую очередь реляционных, разбор примеров запросов на языке SQL и решения задач. Для успешного освоения курса слушатель должен иметь базовые знания математической логики и основ программирования.

1. Цели и задачи

Цель дисциплины

Курс «Базы Данных» рассчитан на студентов, владеющих основами программирования, и предполагает знание базовых принципов работы компьютера - работы с памятью и дисковой подсистемой. Студенты знакомятся с основами реляционной алгебры, языком SQL, знакомятся с общим устройством СУБД, учатся проектировать схему базы данных для решения прикладной задачи, изучают принципы работы оптимизатора запросов, знакомятся с механизмами обеспечения отказоустойчивости и корректного конкурентного доступа.

Задачи дисциплины

- ознакомление слушателей с задачами, требующими для использования базы данных;
- изучение существующих реляционных БД;
- приобретение слушателями навыка использования SQL-запросов;
- умение проектировать базы данных.

2. Перечень формируемых компетенций

Освоение дисциплины направлено на формирование следующих компетенций:

Код и наименование компетенции	Индикаторы достижения компетенции
ОПК-3 Способен использовать и адаптировать существующие математические методы и системы программирования для разработки и реализации алгоритмов решения прикладных задач	ОПК-3.1 Применяет знания основных математических методов и владеет навыками их адаптации для решения конкретной прикладной задачи
	ОПК-3.2 Разрабатывает и реализует алгоритм решения прикладной задачи
	ОПК-3.3 Демонстрирует практический опыт решения прикладных задач с использованием систем программирования и специализированного программного обеспечения

3. Перечень планируемых результатов обучения по дисциплине (модулю)

В результате освоения дисциплины обучающиеся должны знать:

- основы реляционной алгебры;
- принципы проектирования баз данных;
- определения нормальных форм;
- общее устройство БД;
- основы SQL;
- основные принципы работы оптимизатора запросов;
- алгоритмы обеспечения отказоустойчивости;
- уровни изоляции;
- принципы работы блокировочного и многоверсионного планировщика.

уметь:

- проектировать БД с использованием ER диаграмм;
- писать эффективные SQL запросы;
- использовать расширенные возможности языка SQL: хранимые процедуры и функции, триггеры;
- создавать транзакции с учетом параллельного выполнения;
- определять и устранять причины мертвых блокировок (deadlock).

владеть:

- инструментарием для работы с БД;
- инструментарием для проектирования БД.

4. Содержание дисциплины (модуля), структурированное по темам (разделам) с указанием отведенного на них количества академических часов и видов учебных занятий

4.1. Разделы дисциплины (модуля) и трудоемкости по видам учебных занятий

№	Тема (раздел) дисциплины	Трудоемкость по видам учебных занятий, включая самостоятельную работу, час.			
		Лекции	Семинары	Лаборат. работы	Самост. работа
1	Основные понятия.	2	2		8
2	Язык SQL.	2	2		8
3	Трехзначная логика.	2	2		8
4	Функциональные зависимости.	4	4		9
5	Конструкции.	2	2		8
6	Операции	4	4		9
7	Быстродействие.	4	4		9
8	Администрирование.	4	4		9
9	Дополнительные возможности.	4	4		8
10	Современные реляционные СУБД.	2	2		8
Итого часов		30	30		84
Общая трудоёмкость		144 час., 4 зач.ед.			

4.2. Содержание дисциплины (модуля), структурированное по темам (разделам)

Семестр: 4 (Весенний)

1. Основные понятия.

Понятие базы данных. Реляционная модель данных. Типы данных и домены. Отношения. Современные реляционные СУБД

2. Язык SQL.

Обзор языка SQL. Конструкция SELECT. Группировка и агрегатные функции.

3. Трехзначная логика.

Трехзначная логика. NULL-значения. Предикаты.

4. Функциональные зависимости.

Целостность данных. Первая, вторая, третья нормальные формы. Ключи. Нормализация баз данных: теория и практика.

5. Конструкции.

Конструкции UPDATE, INSERT, DELETE.

6. Операции

Блокировки. Транзакции. Требования ACID. Уровни изоляции. Причины возникновения deadlocks и методы борьбы с ними.

7. Быстродействие.

Быстродействие запросов. Методы оптимизации. Физическое устройство реляционной базы данных.

8. Администрирование.

Администрирование баз данных. Роль DBA. Обеспечение отказоустойчивости и катастрофоустойчивости. Этапы сертификации. DDL.

9. Дополнительные возможности.

Дополнительные возможности языка T-SQL. Табличные и скалярные пользовательские функции. Пользовательские процедуры. Работа с метаданными.

10. Современные реляционные СУБД.

Понятие базы данных. Реляционная модель данных. Типы данных и домены. Отношения. Современные реляционные СУБД.

5. Описание материально-технической базы, необходимой для осуществления образовательного процесса по дисциплине (модулю)

Учебная аудитория, оснащенная компьютером и мультимедийным оборудованием (проектор, звуковая система).

6. Перечень рекомендуемой литературы

Основная литература

Дополнительная литература

7. Перечень ресурсов информационно-телекоммуникационной сети "Интернет", необходимых для освоения дисциплины (модуля)

<http://www.sqlpass.org/>
<http://www.sql.ru>
<http://citforum.ru/database/osbd/contents.shtml>
<https://www.sql-ex.ru/>
<https://www.postgresql.org/files/documentation/pdf/11/postgresql-11-A4.pdf>
<https://postgrespro.ru/docs/postgresql>
https://www.dalibo.org/_media/understanding_explain.pdf
https://wiki.postgresql.org/images/4/45/Explaining_EXPLAIN.pdf

8. Перечень информационных технологий, используемых при осуществлении образовательного процесса по дисциплине (модулю), включая перечень необходимого программного обеспечения и информационных справочных систем (при необходимости)

Необходимое программное обеспечение:

- СУБД PostgreSQL не ниже 9.5 (<https://www.postgresql.org/download/>)
- На усмотрение учащегося:
- <https://www.gliffy.com/>
- <https://www.lucidchart.com/pages/>
- Опционально: <https://www.jetbrains.com/datagrip/>

9. Методические указания для обучающихся по освоению дисциплины (модуля)

Формат курса предполагает получение студентом уверенных практических навыков, подкрепленных фундаментальными знаниями. С этой целью студент должен усвоить основные понятия и определения, большая часть которых дается в начале курса. Решение всех примеров и задач студент должен самостоятельно прорабатывать.

Многие вопросы, которые лишь поверхностно затрагиваются на лекциях (или не затрагиваются вовсе), но, тем не менее, являются полезными для более глубокого понимания предмета, можно изучить по основной и дополнительной литературе. Также рекомендуется прорабатывать источники, предлагаемые преподавателем на аудиторных занятиях.

Для получения дополнительных практических навыков и более глубокого освоения основного содержания курса рекомендуется сделать следующие два упражнения:

– Выполнить программную реализацию основных операций реляционной алгебры. Это даст представление об объеме вычислений, которые реальная СУБД выполняется в процессе обработки запросов.

– Выполнить реализацию на каком-либо языке программирования общего назначения (C#, C++, Java) небольшой программы-оболочки, позволяющей просматривать перечень таблиц в данной БД и выполнять с каждой из них CRUD-операции. Это позволит получить практические навыки программного взаимодействия с СУБД.

Предусмотренная по курсу семестровая контрольно-тестовая работа рассчитана на проверку как теоретических знаний, проверка практических знаний производится на контрольных семинарских контрольных работах, подготовленных семинаристом.

Зачет выставляется по итогам работы в семестре, результатам семестровой контрольной, а также устной беседы с преподавателем семинарских занятий.

Литература для самостоятельного изучения:

1. A. Oppel, R. Sheldon SQL: A beginner's guide – 2009, McGill's university, 533 p,
2. SQL для профессионалов /Джо Селко – М.: Лори, 464с. 2009.
3. M. Hernandez Database desing for mere mortals – Pearson Tech, 2003.
4. Next Generation Databases NoSQL, NewSQL, and Big Data – Harrison G.

ОЦЕНОЧНЫЕ МАТЕРИАЛЫ ПО ДИСЦИПЛИНЕ (МОДУЛЮ)

по направлению: Прикладная математика и информатика
профиль подготовки: Искусственный интеллект и большие данные
Сетевое обучение
кафедра алгоритмов и технологий программирования
курс: 2
квалификация: бакалавр

Семестр, формы промежуточной аттестации: 4 (весенний) - Экзамен

Разработчик: Н.А. Меркурьева, ассистент

1. Компетенции, формируемые в процессе изучения дисциплины

Код и наименование компетенции	Индикаторы достижения компетенции
ОПК-3 Способен использовать и адаптировать существующие математические методы и системы программирования для разработки и реализации алгоритмов решения прикладных задач	ОПК-3.1 Применяет знания основных математических методов и владеет навыками их адаптации для решения конкретной прикладной задачи
	ОПК-3.2 Разрабатывает и реализует алгоритм решения прикладной задачи
	ОПК-3.3 Демонстрирует практический опыт решения прикладных задач с использованием систем программирования и специализированного программного обеспечения

2. Показатели оценивания компетенций

В результате изучения дисциплины «Базы данных» обучающийся должен:

знать:

- основы реляционной алгебры;
- принципы проектирования баз данных;
- определения нормальных форм;
- общее устройство БД;
- основы SQL;
- основные принципы работы оптимизатора запросов;
- алгоритмы обеспечения отказоустойчивости;
- уровни изоляции;
- принципы работы блокировочного и многоверсионного планировщика.

уметь:

- проектировать БД с использованием ER диаграмм;
- писать эффективные SQL запросы;
- использовать расширенные возможности языка SQL: хранимые процедуры и функции, триггеры;
- создавать транзакции с учетом параллельного выполнения;
- определять и устранять причины мертвых блокировок (deadlock).

владеть:

- инструментарием для работы с БД;
- инструментарием для проектирования БД.

3. Перечень типовых (примерных) вопросов, заданий, тем для подготовки к текущему контролю

Примеры контрольных заданий:

Типовое задание предполагает написание типового запроса на языках реляционной алгебры / SQL:

1. Запросы на извлечение данных из одного отношения с помощью одной-двух реляционных операций (ограничение, проекция)
2. Запросы на извлечение данных из двух-трех отношений с помощью двух-трех реляционных операций (включая, как минимум, одно соединение)
3. Запросы на извлечение данных из двух-трех отношений с помощью двух-трех реляционных операций (включая, как минимум, одно деление)
4. Запросы на извлечение данных из нескольких отношений с помощью нескольких реляционных операций (включая соединение или/и деление и теоретико-множественные операции)

4. Перечень типовых (примерных) вопросов и тем для проведения промежуточной аттестации обучающихся

Перечень контрольных вопросов:

1. Понятие предметной области. Понятие модели предметной. Понятие модели данных.
2. Сетевая и иерархическая модели.
3. Основные понятия реляционной модели: домен, отношение, кортеж, атрибут, операции над отношениями.
4. Типы данных: скалярные и нескаларные.
5. Понятие базы данных и реляционной базы данных.
6. Понятие переменной отношения (реляционной переменной).
7. Понятие кортежной переменной.
8. Переменная, определенная на домене.
9. Операция соединения: внутреннее, внешнее (левое, правое), полное
10. Операция деления.
11. Эквивалентность реляционной алгебры и реляционного исчисления на кортежах
12. Эквивалентность реляционной алгебры и реляционного исчисления на доменах.
13. Отношение, тип, объект, домен, кортеж: взаимосвязь понятий.
14. Язык SQL: структура запросов на языке SQL.
15. Язык SQL: связь с реляционной алгеброй.
16. Язык SQL: связь с реляционным исчислением на кортежах.
17. Язык SQL. Работа с отсутствующими значениями (NULL).
18. Вложенные запросы в языке SQL.
19. Структура хранимой процедуры/функции в языке SQL.
20. Понятие языка определения данных. Определение пользовательского типа данных.
21. Понятие языка определения данных. Создание таблицы.
22. Понятие языка определения данных. Определение ограничений.
23. Триггеры.
24. Особенности хранимых процедур и функций в СУБД MS SQL Server.
25. Особенности хранимых процедур и функций в СУБД Oracle. Пакеты.
26. Общая архитектура СУБД.
27. Понятие транзакции.
28. Организация хранения данных на жестком диске.
29. Виды (представления, views). Материализованные представления.
30. Индексы: назначение и организация.

Пример билетов:

Билет 1

1. Что такое база данных и модель данных. Иерархическая и сетевая модели данных: свойства, сходства и различия, достоинства и недостатки.
2. Хранилища данных. Принципы создания, модели данных.
3. Даны отношения:
 - a. PRODUCT (product_id, product_nm, product_price, manufacturer_id, product_type)
 - b. MANUFACTURER (manufacturer_id, manufacturer_nm, phone_no)
 - c. SHOP (shop_id, shop_nm, city_nm, street_nm, building_no)
 - d. SUPPLY (supply_id, shop_id, product_id, supply_dt, item_cnt, supplier_nm) – Факт поставкиДля каждого поставщика рассчитать процент товаров каждого производителя в поставках за последний месяц. Пронумеровать производителей в рамках одного поставщика в порядке возрастания рассчитанного значения.

Билет 2

1. SQL: DML. Описать синтаксис SELECT-запросов, порядок выполнения. Связь реляционной алгебры с SQL.
2. Хранимые процедуры и функции.
3. Построить логическую модель базы данных для футбола. Должна быть возможность получить информацию о голах: кто забил, на какой минуте, кто отдал голевую. Также необходимо хранить информацию обо всех трансферах.

Критерии оценивания

- оценка «отлично (10)» выставляется студенту, показавшему всесторонние, систематизированные, глубокие знания учебной программы дисциплины и умение уверенно применять их на практике при решении конкретных задач, свободное и правильное обоснование принятых решений
- оценка «отлично (9)» выставляется студенту, показавшему всесторонние, систематизированные, глубокие знания учебной программы дисциплины и умение применять их на практике при решении конкретных задач, свободное и правильное обоснование принятых решений
- оценка «отлично (8)» выставляется студенту, показавшему всесторонние систематизированные, глубокие знания учебной программы дисциплины и умение применять их на практике при решении конкретных задач, и правильное обоснование принятых решений
- оценка «хорошо (7)» выставляется студенту, если он твердо знает материал, грамотно и по существу излагает его, умеет применять полученные знания на практике, но допускает в ответе или в решении задач некоторые неточности;
- оценка «хорошо (6)» выставляется студенту, если он знает материал, грамотно и по существу излагает его, умеет применять полученные знания на практике, но допускает в ответе или в решении задач некоторые неточности;
- оценка «хорошо (5)» выставляется студенту, если он знает материал, и по существу излагает его, умеет применять полученные знания на практике, но допускает в ответе или в решении задач некоторые неточности;
- оценка «удовлетворительно (4)» выставляется студенту, показавшему фрагментарный, разрозненный характер знаний, недостаточно правильные формулировки базовых понятий, нарушения логической последовательности в изложении программного материала, но при этом он владеет основными разделами учебной программы, необходимыми для дальнейшего обучения и может применять полученные знания по образцу в стандартной ситуации;
- оценка «удовлетворительно (3)» выставляется студенту, показавшему фрагментарный, разрозненный характер знаний, недостаточно правильные формулировки базовых понятий, нарушения логической последовательности в изложении программного материала, но при этом он владеет фрагментарно основными разделами учебной программы, необходимыми для дальнейшего обучения и может применять полученные знания по образцу в стандартной ситуации;
- оценка «неудовлетворительно (2)» выставляется студенту, который не знает большей части основного содержания учебной программы дисциплины, допускает грубые ошибки в формулировках основных понятий дисциплины и не умеет использовать полученные знания при решении типовых практических задач
- оценка «неудовлетворительно (1)» выставляется студенту, который не знает формулировок основных понятий дисциплины.

5. Методические материалы, определяющие процедуры оценивания знаний, умений, навыков и (или) опыта деятельности

Дифференцированный зачет проводится с учетом текущей успеваемости и результатов сдачи курсовой работы. При необходимости, в процессе собеседования со студентом проводится выборочный опрос на знание контрольных вопросов, предлагаются типовые задачи.