

Лекция №1

# Базы данных

Моисеев Михаил Юрьевич

Вводная лекция

# Цели и задачи курса

## Получение навыков проектирования ИС

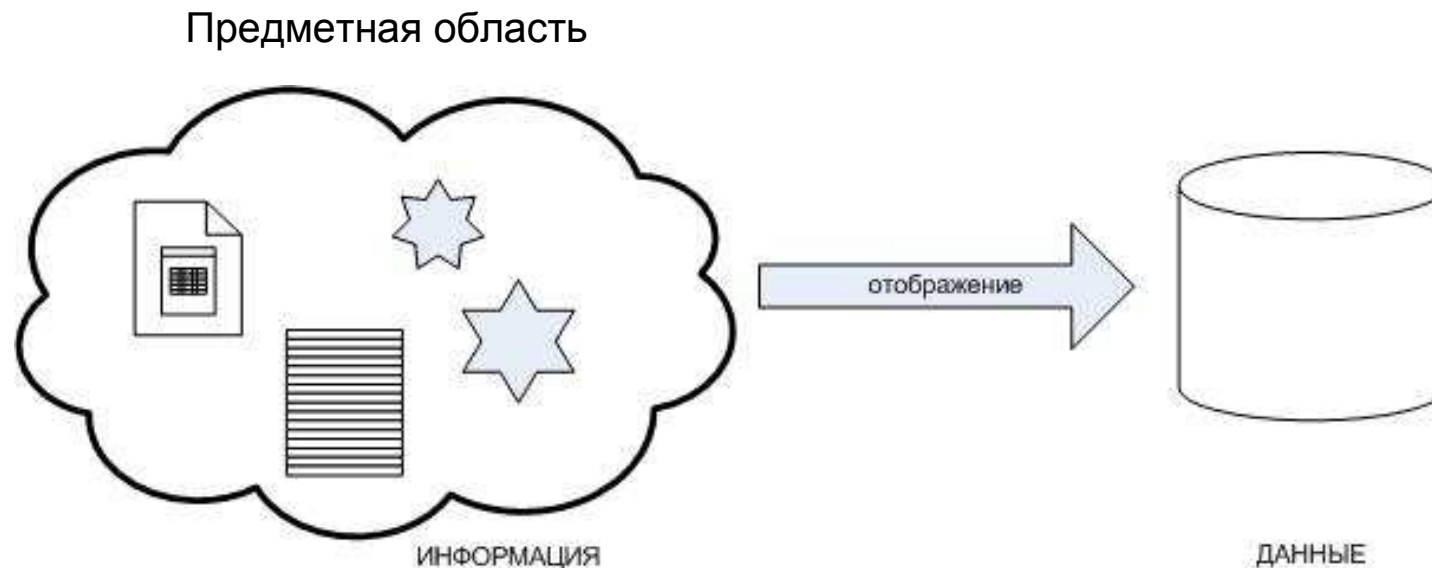
- Основные понятия и определения
- Базовые теории
- Практические навыки проектирования ИС
- Изучение современных технологий и средств разработки

# Организация лекций

- Общие правила
- Материалы находятся на [tiger.ftk.spbstu.ru/trac](http://tiger.ftk.spbstu.ru/trac)
- Ваши вопросы в течении лекции
- Мои вопросы в конце лекции
- Презентации выложены **не будут**

# Информация и данные

- **Информация** - любые сведения о каком-либо событии
- **Данные** - информация, фиксированная в определенной форме, удобной для хранения и обработки



Вариантность отображения информации

# Информационная система

- **ИС** обеспечивает **динамическую** информационную модель предметной области и представляет ее в удобной форме
- **Инфологический** аспект
  - Какие объекты реального мира будут приняты в рассмотрение
  - Какие характеристики и взаимосвязи этих объектов будут учитываться
- **Даталогический** аспект
  - Разрабатываются формы представления информации
  - Приводятся модели и методы преобразования данных

# Классификация ИС

## По масштабу

- Индивидуальные
- Групповые
- Корпоративные
- Глобальные

## По области применения

- Информационно-справочные, поисковые
- Управление контентом
- Управление деятельностью (предприятием)
- ГИС

## По способу распределения

- Локальные
- Распределенные

# Классификация ИС #2

## По способу построения

- Файловые
- Клиент-серверные
- Трех и более уровневые

## По используемым моделям данных

- Реляционные
- Объектно-ориентированные (объектно-реляционные)
- Сетевые
- Иерархические
- Другие модели

# История развития ИС

- Файловые ИС
- Появление первой БД – 196х год
- Первое поколение БД - иерархические и сетевые
- Второе поколение БД – реляционные
- Третье поколение БД – ООБД и ОРБД



# Файловые ИС

- **Файловая ИС** - набор прикладных программ, хранящие данные в файлах в некотором формате
- Недостатки файловых систем
  - Разделение и изоляция данных
  - Дублирование данных
  - Зависимость данных и программ
  - Несовместимость форматов файлов
  - Дублирование кода

# ИС основанные на БД

- **БД** – разделяемый набор логически связанных данных вместе с описаниями данных
- **СУБД** - это СПО, которое дает возможность пользователю управлять данными и структурой БД
- Возможности СУБД
  - Позволяет создавать БД, изменять ее структуру (метаданные)
  - Обеспечивает универсальные средства для работы с данными
  - Контролируемый доступ к БД
    - Контроль целостности
    - Параллельная работа нескольких клиентов
    - Резервное копирование и восстановление после сбоев
    - Защита данных

# ИС основанные на БД #2

- Преимущества использования БД
  - Стандартизованный формат хранения данных
  - Доступ к данным через один интерфейс
  - Доступ к данным осуществляется на более высоком уровне
  - Обеспечивается независимость данных и программ
  - Более эффективные алгоритмы работы с данными
  - Ускоряется и упрощается разработка конечных ИС
  - Гибкость разработанных решений, возможность перехода на другую платформу или другую БД

# ИС основанные на БД #3

- Дополнительные преимущества использования БД
  - Развитые службы резервного копирования и восстановления
  - Контроль за избыточностью данных
  - Автоматическая поддержка целостности данных
  - Улучшенное управление параллельной работой
  - Улучшение показателей производительности
- Недостатки использования БД
  - Дополнительные затраты на преобразование данных
  - Производительность при обработке больших объемов слабоструктурированных данных
  - Дополнительная стоимость

# Уровни абстракции

- СУБД позволяет разделить обработку данных на несколько уровней абстракции
  - Внешнее представление
    - Часть информации БД, настраивается под пользователя
    - Несколько различных представлений
  - Концептуальный уровень
    - Представляет семантическую информацию
    - Использует язык выбранной модели данных
  - Внутренний уровень
    - Описание формата данных с точки зрения их хранения в БД
  - Физический уровень
    - Описание данных с точки зрения хранения на физическом носителе.

# Уровни абстракции #2



# Модели данных

- **Модель данных** – набор понятий, описывающий данные, способы обработки данных и связи между ними, а также ограничения, накладываемые на данные.
- Модель данных должна определить
  - Структуры
  - Ограничения
  - Операции

# Модель сущность-связь

- Модель **сущность-связь** (Entity-Relationship, **ER**) была предложена Питером Ченом в 1976 г
- Основана на семантической информации о предметной области – концептуальный уровень
- Определяет значения данных в контексте их взаимосвязи с другими данными
- Является наиболее общей – из нее могут быть порождены все существующие модели данных (иерархическая, сетевая, реляционная, объектная)
- Модель сущность-связь не определяет операций над данными и ограничивается описанием только их логической структуры



# Модель сущность-связь #2

- Любой фрагмент предметной области может быть представлен как множество **сущностей**, между которыми определено некоторое множество **связей**
- **Сущность** - объект, который может быть идентифицирован неким способом, отличающим его от других объектов
- **Набор сущностей** - множество сущностей одного типа (с одинаковым набором атрибутов)
- Сущность определяется значениями атрибутов, которые описывают свойства всех членов данного набора сущностей
- Множество значений (область определения) атрибута называется **доменом**

# Модель сущность-связь #3

**Связь** - ассоциация, установленная между несколькими сущностями

Связь реализуется между двумя наборами сущностей

- Главный
- Подчиненный

Подчиненные сущности связаны с (ссылаются на) главной сущностью.

**Степень связи** - число сущностей, которое ассоциировано с помощью данной связи

- Связь 1 : 1
- Связь 1 : N
- Связь N : N

Обязательные и необязательные связи

# Модель сущность-связь #4

Предметная область: кафедра КСПТ

Отображение: студенты, группы, специальности, изучаемые предметы, преподаватели, участие студентов в НИР кафедры

Студент: (**Имя, Фамилия, Отчество**, Дата рождения, Номер паспорта, Номер аттестата, Место проживания, **Ссылка на группу, Ссылка на НИР ...**).

Группа: (**Номер группы, Куратор, Ссылка на специальность, ...**).

НИР: (**Название темы**, Руководитель, Заказчик работы, ...).

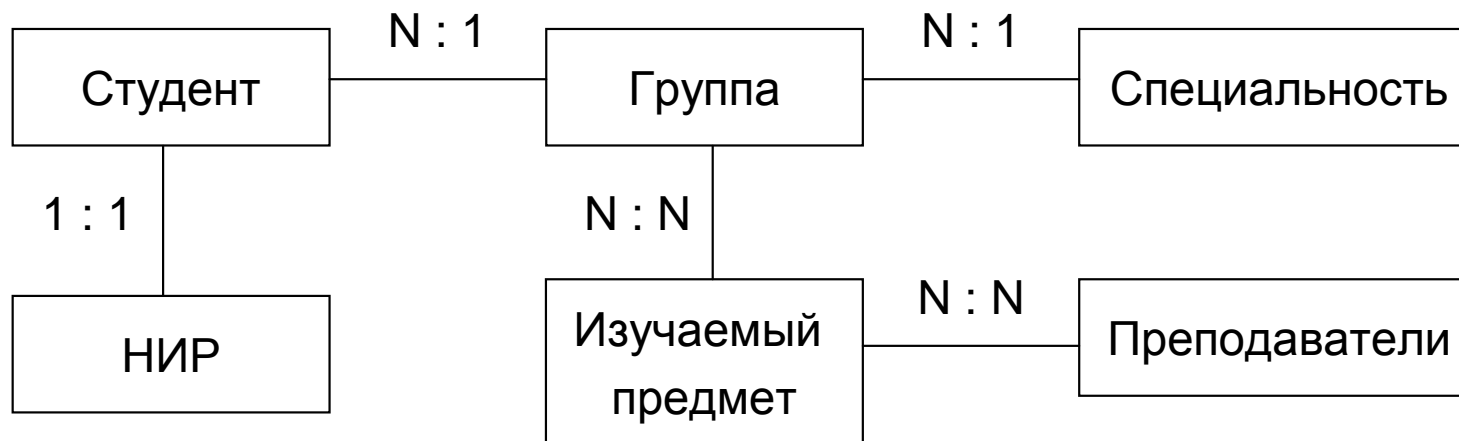
Преподаватели : (...)

Специальности : (...)

Предметы : (...)

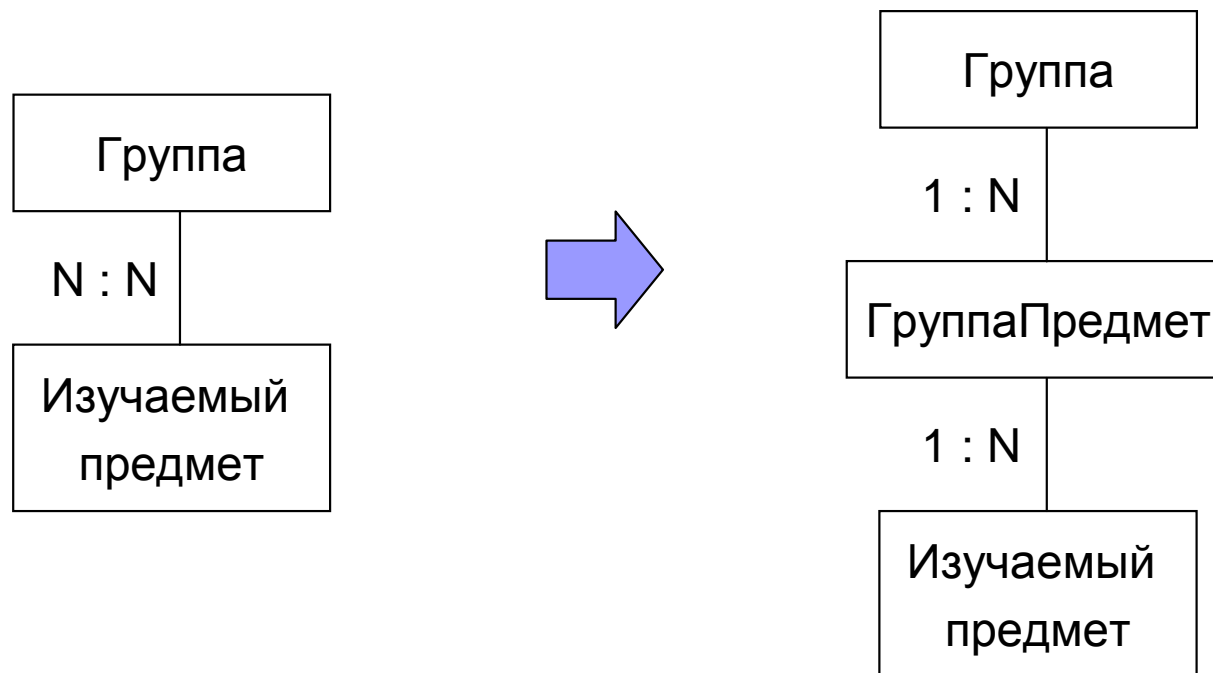
# Модель сущность-связь #5

Представлены все сущности и связи между ними



# Модель сущность-связь #6

Для реализации связи N:N необходимо добавить новую сущность и связать ее с исходными сущностями (как подчиненную).



ГруппаПредмет: (Ссылка на группу, Ссылка на предмет)

# Вопросы

- Чем информация отличается от данных?
- Можно ли извлечь разную информацию из одной предметной области?
- Какова основная задача ИС?
- Перечислите основные отличия файловых и БД ИС.
- Могут ли в набор сущностей входить сущности с разными атрибутами?
- Между какими сущностями устанавливаются связи?
- Какие виды связей используются в модели сущность-связь?
- Перечислите уровни абстракции данных?
- Проведите классификацию ИС Intranet?