

СИСТЕМЫ УПРАВЛЕНИЯ БАЗАМИ ДАННЫХ (I ЧАСТЬ)

1

Введение в базы данных

КОНТАКТЫ

Медведев Михаил Александрович

Почта – stipleee@gmail.com

Телеграм - [@MedvedevMA](https://www.instagram.com/MedvedevMA)

ЦЕЛЬ. ПРЕДМЕТ ИЗУЧЕНИЯ

- ▣ **Цель.** Раскрыть принципы построения систем управления базами данных (СУБД); ознакомить слушателей с исследованиями в области моделирования СУБД; помочь освоить навыки моделирования баз данных (БД).
- ▣ **Предмет изучения.** Базовые понятия и определения. Назначение и структура СУБД. Основные задачи и методы БД и их компонентов.

О КУРСЕ

- Лекции – 7 занятий;
- Лабораторные – 5 занятий;
- Зачет – 2 занятия.

ТЕМЫ ЛЕКЦИОННЫХ ЗАНЯТИЙ

- Введение в базы данных;
- Проектирование баз данных;
- Реляционная модель;
- Нормализация отношений;
- Языковые средства СУБД. Язык SQL;
- Защита данных. Администрирование БД;
- Распределенные БД. Хранилища данных. Обзор современных СУБД и средств автоматизированного проектирования БД.

РЕКОМЕНДУЕМАЯ ЛИТЕРАТУРА

- К. Дж. Дейт. Введение в системы баз данных.
- Головчинер М.Н. Базы данных. Основные понятия. Модели данных. Процесс проектирования
- Бен Форта SQL. 10 Минут на урок.

ВВЕДЕНИЕ В БАЗЫ ДАННЫХ

1. ПРЕДМЕТНАЯ ОБЛАСТЬ, ИНФОРМАЦИЯ И ДАННЫЕ

- Предметная область — часть реального мира, рассматриваемая в пределах контекста;
- Информация – сведения, знания, сообщения;
- Данные - сведения, факты, показатели, выраженные в какой-либо форме (формате);
- Данные - это информация, представленная в пригодном для ее обработки виде.

СУЩНОСТЬ, АТТРИБУТ, ЗАПИСЬ

- Сущность – объект из предметной области;
- Однородные сущности – объекты, обладающие одинаковым набором свойств;
- Тип сущности - множество однородных объектов, понимаемое как единое целое;
- Атрибут – это информационное отображение свойства сущности;
- Запись – совокупность атрибутов сущности.

ПРИМЕРЫ АТТРИБУТОВ

- Работник фирмы: ФИО, табельный номер, стаж, уровень допуска,...
- Автомобиль: марка, модель, объем двигателя,...
- Денежный перевод: отправитель, сумма, статус,...
- Кассовый чек: время операции, сумма скидки, ФИО оператора,...

ДАННЫЕ И ИНТЕРПРЕТАЦИЯ ДАННЫХ

интерпретация

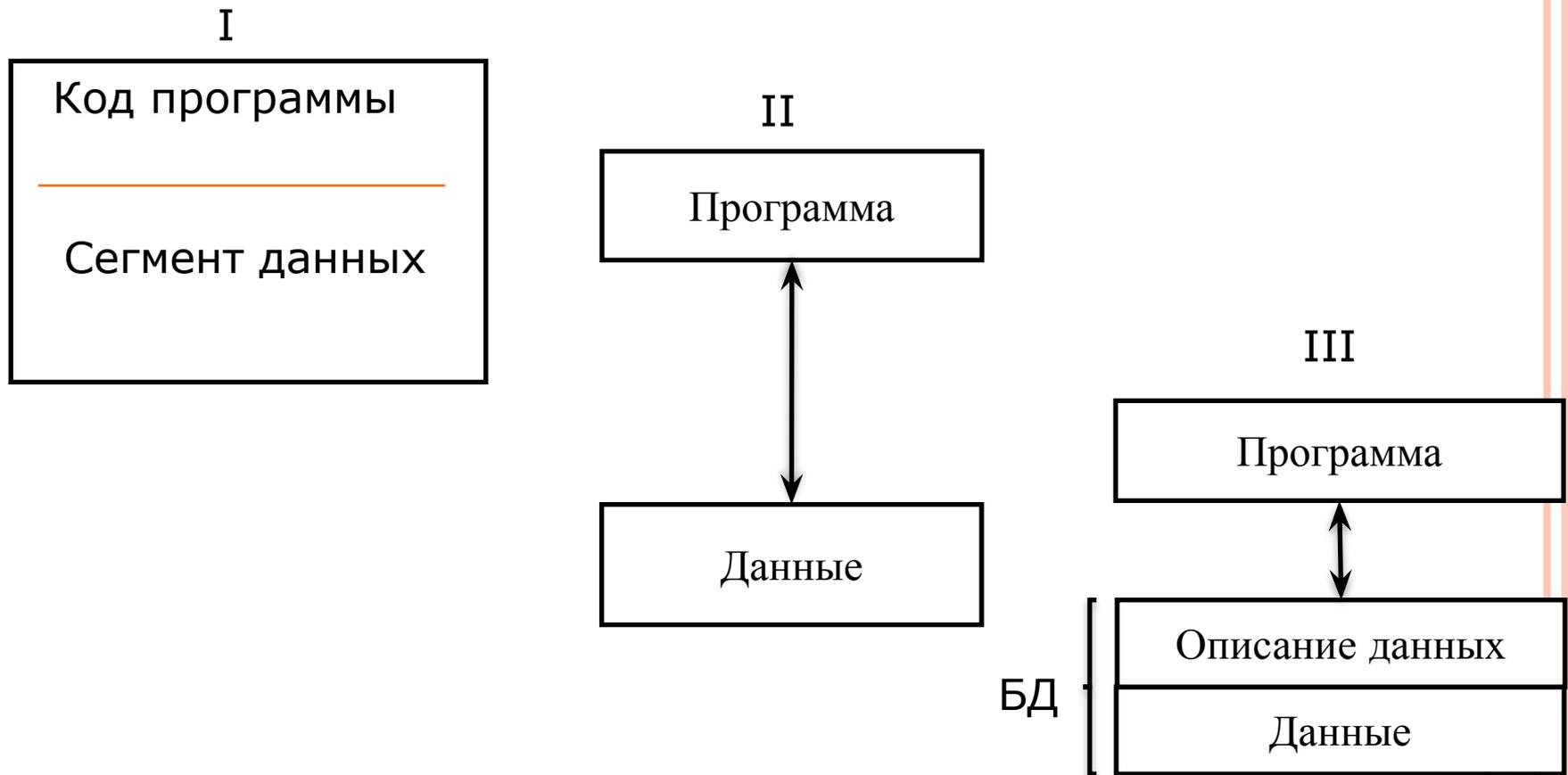
Номер рейса	Дни недели	Пункт отправления	Время вылета	Пункт назначения	Время прибытия	Тип самолета	Стоимость билета
138	2_4_7	Баку	21.12	Москва	0.52	ИЛ-86	115.00
57	3_6	Ереван	7.20	Киев	9.25	ТУ-154	92.00
1234	2_6	Казань	22.40	Баку	23.50	ТУ-134	73.50
242	1 по 7	Киев	14.10	Москва	16.15	ТУ-154	57.00
86	2_3_5	Минск	10.50	Сочи	13.06	ИЛ-86	78.50
137	1_3_6	Москва	15.17	Баку	18.44	ИЛ-86	115.00
241	1 по 7	Москва	9.05	Киев	11.05	ТУ-154	57.00
577	1_3_5	Рига	21.53	Таллин	22.57	АН-24	21.50
78	3_6	Сочи	18.25	Баку	20.12	ТУ-134	44.00
578	2_4_6	Таллин	6.30	Рига	7.37	АН-24	21.50

данные

ПРИЧИНЫ РАЗДЕЛЕНИЯ ДАННЫХ И ИХ ИНТЕРПРЕТАЦИИ ПРИ РАБОТЕ НА ЭВМ

- ЭВМ не обладали достаточными возможностями для интерпретации данных на естественном языке – основном языке интерпретации данных;
- Стоимость памяти ЭВМ первоначально была весьма высока, и потому память использовалась только для хранения самих данных, а интерпретация данных возлагалась на программиста-пользователя;
- Искусственно создавалась жесткая зависимость между данными и интерпретирующими их программами;
- При использовании одних и тех же данных разными программами возникает масса проблем

ЭВОЛЮЦИЯ ПРИНЦИПОВ ОБРАБОТКИ ДАННЫХ



2. ПОНЯТИЕ ИНФОРМАЦИОННОЙ СИСТЕМЫ

- Обработка информации – основа решения многих задач;
- Информационные системы (ИС) создаются для облегчения обработки информации;
- Автоматизированные ИС – ИС, в которых применяются технические средства (ЭВМ);
- Большинство ИС являются автоматизированными.

- ИС – любая система обработки информации.
- ИС – совокупность аппаратно-программных средств, задействованных для решения прикладной задачи.

КЛАССИФИКАЦИЯ ИС ПО ОБЛАСТИ ПРИМЕНЕНИЯ

- Производство;
- Образование;
- здравоохранение;
- Наука;
- Социальная сфера;
- Торговля;
- ...

КЛАССИФИКАЦИЯ ИС ПО ЦЕЛЕВОЙ ФУНКЦИИ

- Управляющие;
- Информационно-справочные;
- Поддержки принятия решений.

3. ОПРЕДЕЛЕНИЕ И НАЗНАЧЕНИЕ БАЗ ДАННЫХ

- Единая концепция по хранению и обработке данных появилась в 60-е годы XX века .
- База данных(БД) – совокупность специальным образом организованных данных, хранимых в памяти вычислительной системы и отображающих состояние объектов и их взаимосвязей в рассматриваемой предметной области.
- База данных (БД) - это совместно используемый интегрированный набор логически связанных данных, хранящийся вместе с описанием этих данных, предназначенный для удовлетворения информационных потребностей организации.

ПРЕИМУЩЕСТВА БД

- Сокращение избыточности в хранимых данных ;
- Возможность контроля целостности и непротиворечивости данных;
- Обеспечение безопасности данных;
- Обеспечение независимости приложений от физической организации данных.

4. СОСТАВ СИСТЕМ БД

- Данные;
- Аппаратное обеспечение;
- Программное обеспечение;
- Люди:
 - Администраторы БД;
 - Администраторы данных;
 - Разработчики БД;
 - Прикладные программисты;
 - Конечные пользователи.

КОМПОНЕНТЫ СИСТЕМЫ БАЗЫ ДАННЫХ



АППАРАТНОЕ ОБЕСПЕЧЕНИЕ БД

- Сервер(-а), CPU, память и т.д.
- СХД: DAS, SAN, NAS
- Внешний доступ: сеть

ПЕРСОНАЛ, РАБОТАЮЩИЙ С БД

- Администраторы БД;
- Администраторы данных;
- Разработчики БД;
- Прикладные программисты;
- Конечные пользователи.

5.ОПРЕДЕЛЕНИЕ СИСТЕМЫ УПРАВЛЕНИЯ БАЗАМИ ДАННЫХ

- Система управления базами данных (СУБД, DBMS) - это программное обеспечение, с помощью которого пользователи могут определять, создавать и поддерживать базу данных, а также осуществлять к ней контролируемый доступ.

- Ключевые возможности СУБД
 - позволяет определять базу данных путем описания типов данных, их структуры и ограничений для хранимой информации на языке определения данных;
 - позволяет добавлять, обновлять, удалять и извлекать информацию с помощью языка управления данными;
 - предоставляет контролируемый доступ к базе данных с помощью:
 - системы защиты от несанкционированного доступа;
 - системы поддержки целостности данных, обеспечивающей их непротиворечивость в ходе хранения;
 - системы поддержки параллельного многопользовательского доступа к базе данных;
 - системы восстановления после сбоев;
 - доступного пользователям каталога с метаданными.

ОСНОВНЫЕ ФУНКЦИИ СУБД

- Непосредственное управление данными во внешней памяти;
- Управление буферами оперативной памяти;
- Управление транзакциями;
- Журнализация и восстановление после сбоев;
- Поддержка языков БД.

ОСНОВНЫЕ ИНСТРУМЕНТАРИИ СУБД

- Модули описания структуры базы данных и средства конструирования экранных форм для ввода, просмотра и обработки данных в диалоговом режиме;
- Инструменты создания стандартных запросов для выборки данных при заданных условиях, а также выполнения операций по их обработке и созданию необходимых отчетов из базы данных с выводом их на печать;
- Группа языковых средств, представленная макросами, встроенными алгоритмическими языками (Dbase, VBa), языком запросов QBE (Query By Example, SQL);
- Средства создания приложений пользователя (генераторы приложений, модули разработки меню и панелей управления приложениями).

ТРЕБОВАНИЯ К СУБД

- Недопущение дублирования и противоречивости данных в различных объектах информационной модели;
- Обеспечение целостности базы данных;
- Возможности модификации ее структуры без повторной загрузки данных;
- Возможность реорганизации размещения данных базы на носителях;
- Возможность многоаспектного доступа к данным для использования их в решаемых задачах;
- Защита и восстановление данных при аппаратных и программных сбоях;
- Защита данных от несанкционированного доступа;
- Наличие языка запросов высокого уровня, ориентированного на конечного пользователя.

ПИОНЕРЫ СУБД

- IMS. IBM, 1968;
- ADABAS. Software AG, 1969;
- IDMS. Cullinet, 1971;
- ИНЭС. ВНИИСИ АН СССР, 1976.

КЛАССИФИКАЦИЯ СУБД

По модели данных



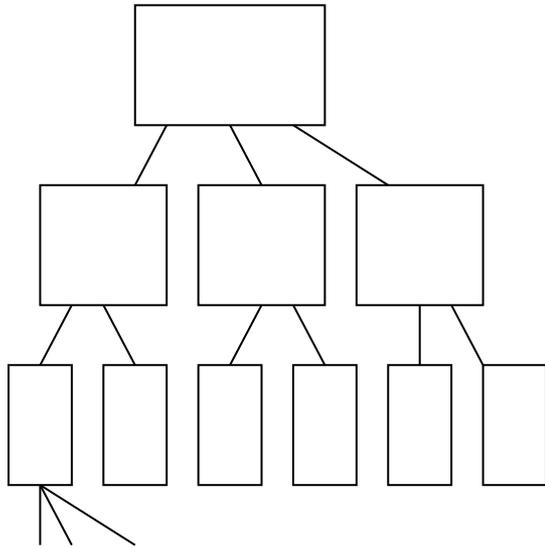
По архитектуре организации хранения данных



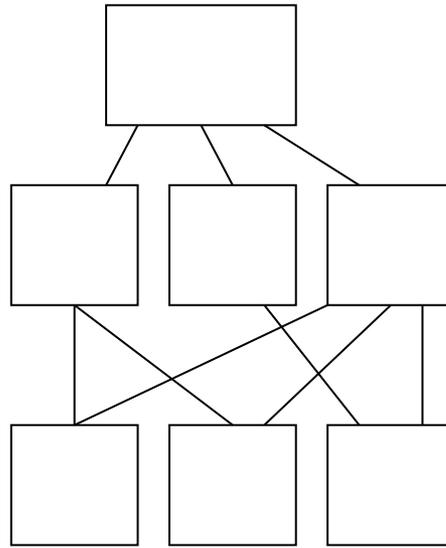
По способу доступа к БД



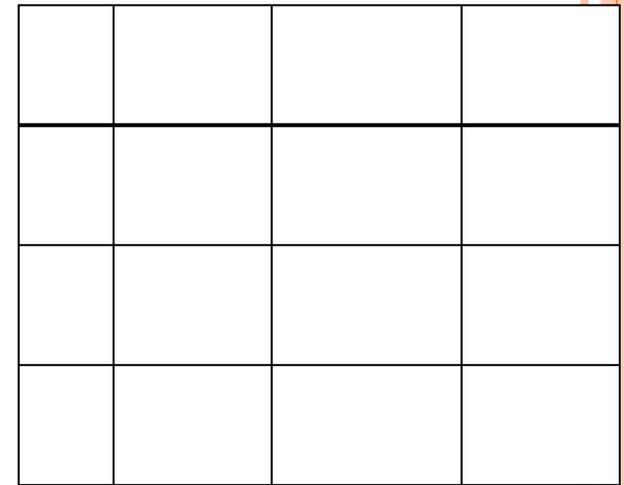
Типы моделей баз данных



Иерархическая



Сетевая



Табличная

6. ЭТАПЫ РАЗРАБОТКИ БД

- Выбор СУБД (определяется функциональными возможностями СУБД, техническими средствами компьютера – конфигурация, память, быстродействие);
- Создание внутримашинной БД (задается структура БД на основе проекта немашинной информационной базы и ее документов, осуществляется ввод данных, выполняются другие необходимые действия с информационной базой в соответствии с функциональными возможностями СУБД) – поэтапно;
- После выявления логических взаимосвязей между информационными компонентами, разработки процедур формализации и моделирования данных приступают к созданию средствами СУБД базы данных на машинных носителях.

УРОВНИ ПРЕДСТАВЛЕНИЯ ДАННЫХ

- Внешний уровень - представление БД с точки зрения пользователей. Этот уровень описывает ту часть БД, которая относится к каждому пользователю. Каждый пользователь имеет дело с представлением БД, выраженным в наиболее удобной для него форме, например, только с теми сущностями, атрибутами и связями, которые для него интересны или с данными в наиболее подходящем формате. Это инфологический уровень;
- Даталогический уровень – описание БД на языке конкретной СУБД, описание собственно данных;
- Физический уровень – физическая организация данных в базе.

МОДЕЛИ ДАННЫХ

- Инфологическая модель отображает реальный мир в некоторые понятные человеку концепции, полностью независимые от параметров среды хранения данных;
- Инфологическая модель должна быть отображена в компьютерно-ориентированную даталогическую модель, "понятную" СУБД;
- Физическая организация данных оказывает основное влияние на эксплуатационные характеристики БД. Разработчики СУБД пытаются создать наиболее производительные физические модели данных, предлагая пользователям тот или иной инструментарий для поднастройки модели под конкретную БД.

МОДЕЛИ ДАННЫХ



ИНФОЛОГИЧЕСКАЯ МОДЕЛЬ

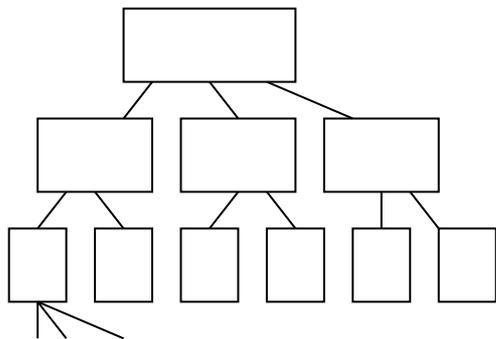
- Инфологическая модель – отображает реальный мир в некоторые понятные человеку концепции, полностью независимые от параметров среды хранения данных;
- Цель инфологического моделирования – обеспечение наиболее естественных для человека способов сбора и представления той информации, которую предполагается хранить в создаваемой БД;
- Основные конструктивные элементы инфологических моделей – СУЩНОСТИ, СВЯЗИ МЕЖДУ НИМИ и АТТРИБУТЫ СУЩНОСТЕЙ;
- Такая инфологическая модель называется моделью СУЩНОСТЬ – СВЯЗЬ.

ДАТАЛОГИЧЕСКАЯ МОДЕЛЬ

- Даталогическая модель – компьютеро-ориентированное отображение инфологической модели;
- Основные типы даталогических моделей:
 - Иерархическая;
 - Сетевая;
 - Реляционная;
 - Объектно-ориентированная.

7. ИЕРАРХИЧЕСКАЯ МОДЕЛЬ ДАННЫХ

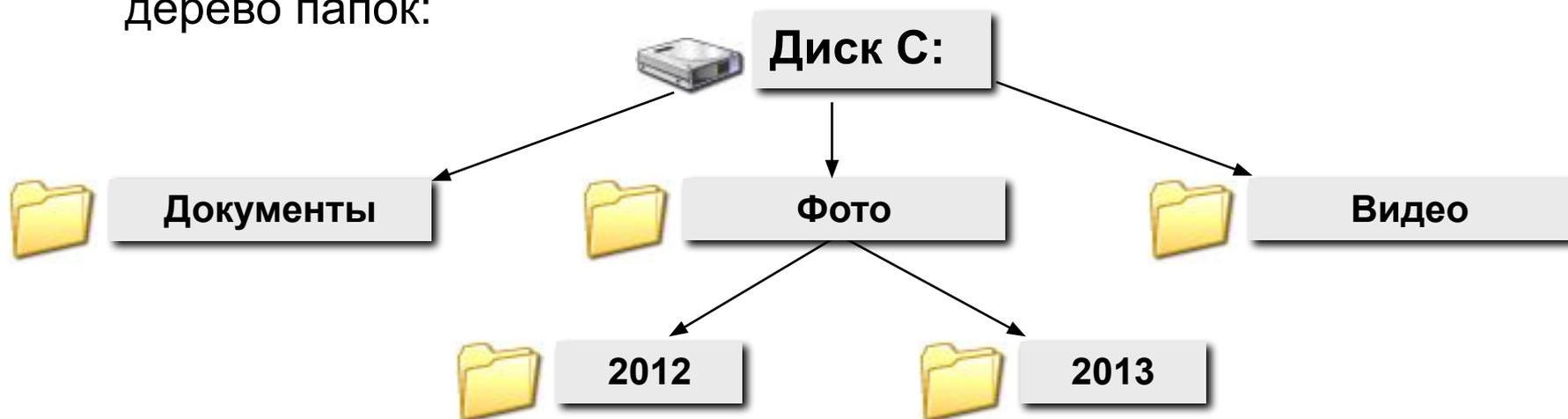
- имеют древовидную структуру, когда каждому узлу структуры соответствует один сегмент, представляющий собой поименованный линейный кортеж данных;
- Каждому сегменту, кроме корневого, соответствует один входной и несколько выходных сегментов;
- Каждый сегмент лежит на единственном иерархическом пути, начинающемся с корневого сегмента. При описании такой логической организации данных достаточно для каждого сегмента указать его входной сегмент. Так как в иерархической модели каждому входному сегменту данных соответствует N выходных, то такие модели весьма удобны для представления отношений типа 1:L в предметной области.



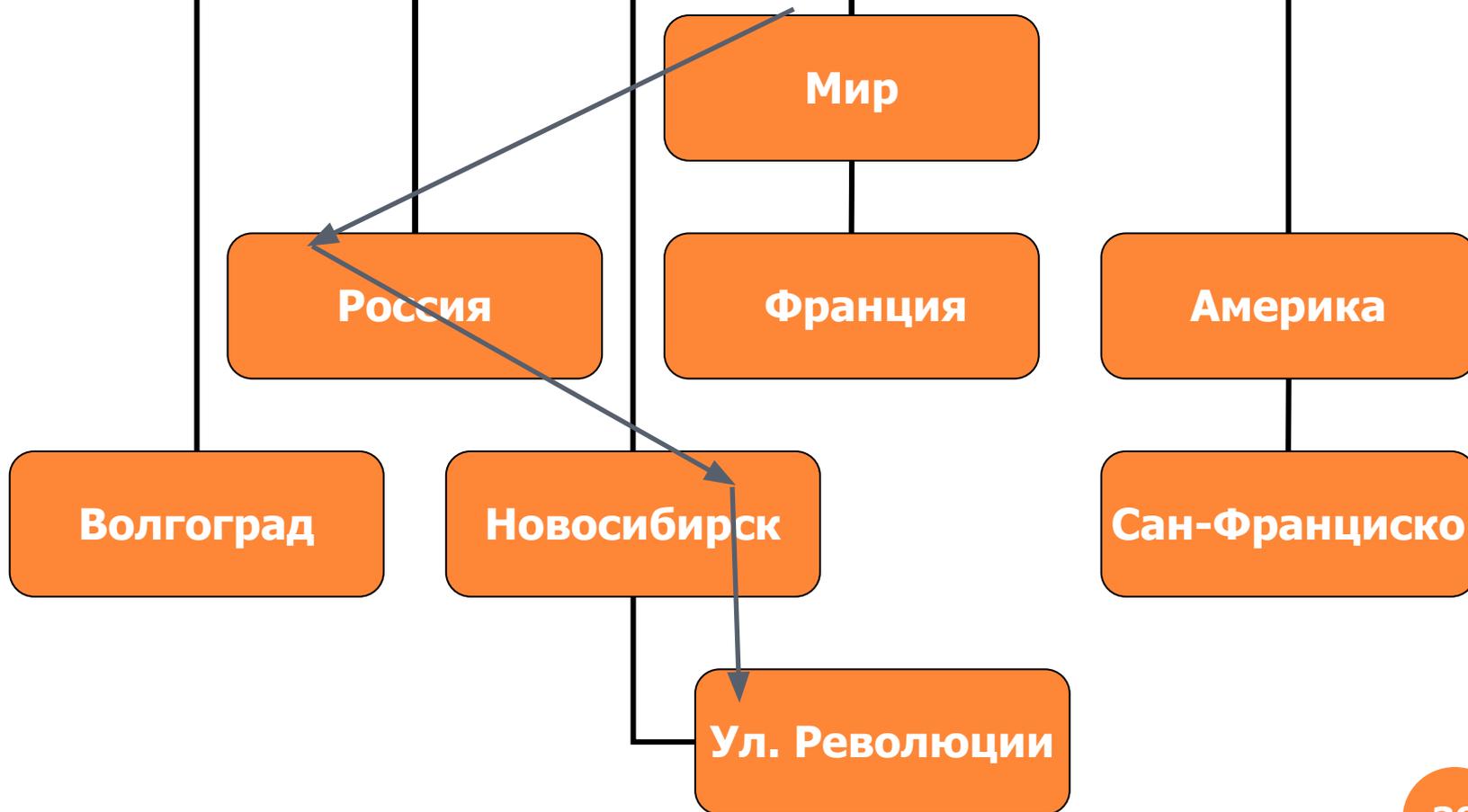
Иерархическая БД – это набор данных в виде многоуровневой структуры (дерева).

УЗЕЛ- информационная модель элемента, находящегося на данном уровне иерархии.

дерево папок:



ИЕРАРХИЧЕСКАЯ МОДЕЛЬ ДАННЫХ



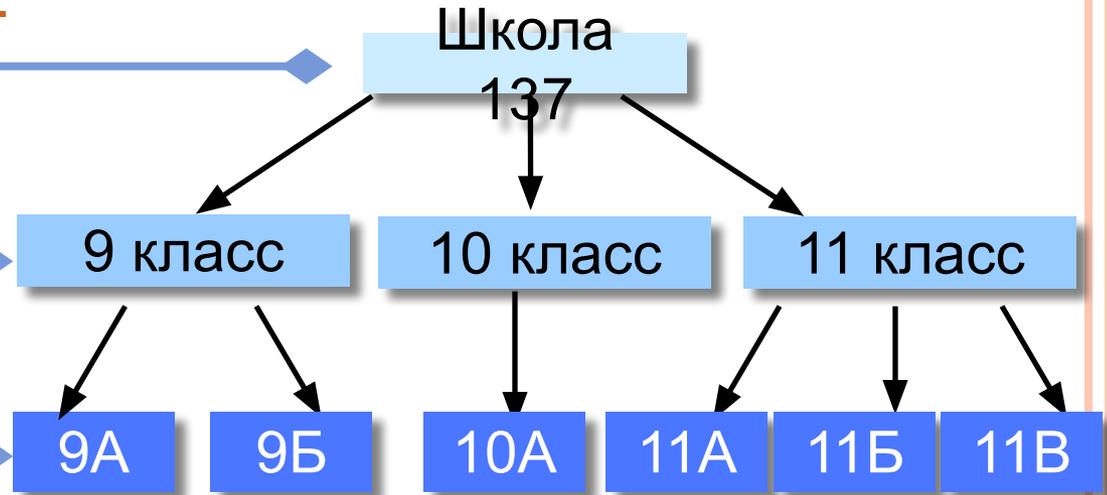
Единственный иерархический путь к сегменту модели, начинающийся с корневого сегмента

Структура школы:

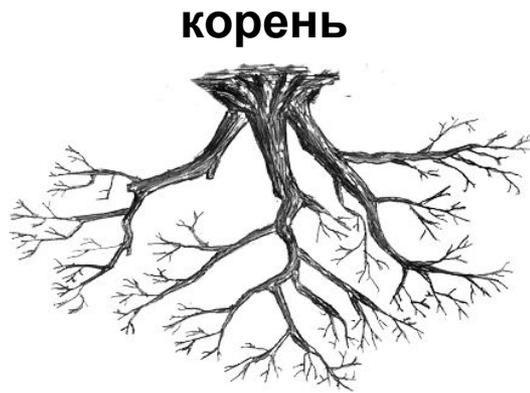
Школа (уровень 1)

Класс (уровень 2)

Параллель (уровень 3)



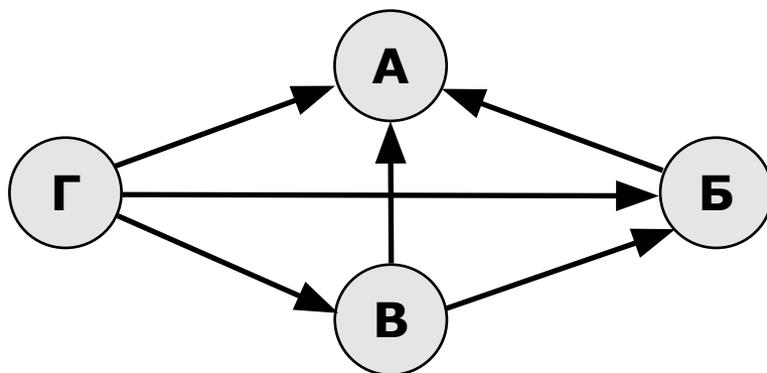
Генеалогическое дерево



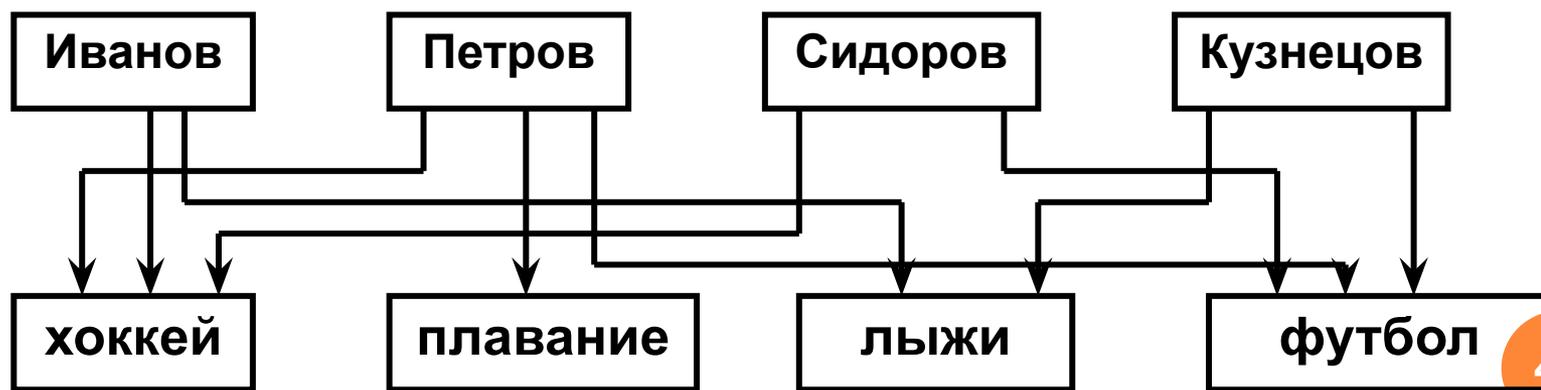
8. СЕТЕВАЯ МОДЕЛЬ ДАННЫХ

- Состоит из «наборов» - двухуровневых деревьев;
- Наборы соединяются с помощью «записей-связок», образуя цепочки;
- Допускается несколько входных элементов наряду с возможностью наличия сегментов без входов;
- Сегменты данных в сетевых базах данных могут иметь множественные связи с сегментами старшего уровня;
- В сетевых моделях данных любая запись старшего уровня может содержать данные, относящиеся к набору записей подчиненного уровня. Обращение к набору всех записей реализуется, начиная с записи старшего уровня.

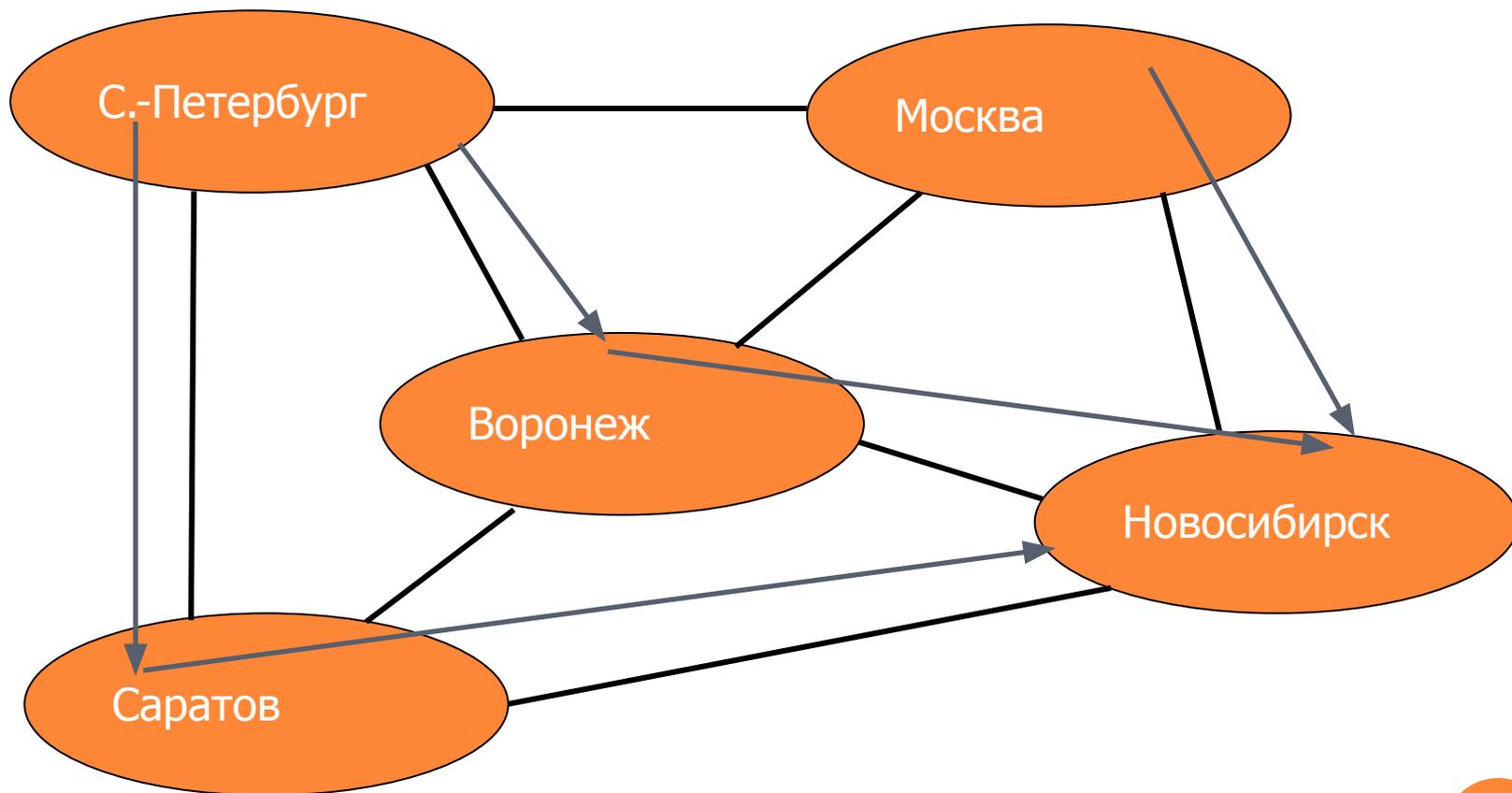
Сетевая модель – это набор узлов, в которых каждый может быть связан с каждым (схема дорог)



Пример: посещение учащимися одной группы спортивных секций



СЕТЕВАЯ МОДЕЛЬ ДАННЫХ



9. РЕЛЯЦИОННАЯ МОДЕЛЬ ДАННЫХ

- В такой модели общая структура данных может быть представлена в виде таблицы, в которой каждая строка значений (кортеж) соответствует логической записи, а заголовки столбцов являются названиями полей записи;
- Первую реляционную модель данных предложил в 1970 году сотрудник фирмы IBM Эдгар Кодд.

ПРЕИМУЩЕСТВА

- Простота структур данных;
- Удобное для пользователя табличным представлением;
- Удобный доступ к данным. Процедуры сохранения и поиска осуществляются с применением операций на множествах (объединение, пересечение, разность, произведение) и реляционных операций (выбрать, спроецировать, соединить, разделить);
- каждому объекту предметной области соответствует одно или более отношений. При необходимости связь между объектами можно указать в явном виде.

РЕЛЯЦИОННАЯ МОДЕЛЬ ДАННЫХ

Таблица «Отдель»

№Отдел	Название
1	Продажи
2	Реклама

Таблица
«Работники»

№Табельный	№Отдел	ФИО	ЗП
10	1	Иванов Иван Иванович	10000
20	2	Петров Петр Петрович	12000

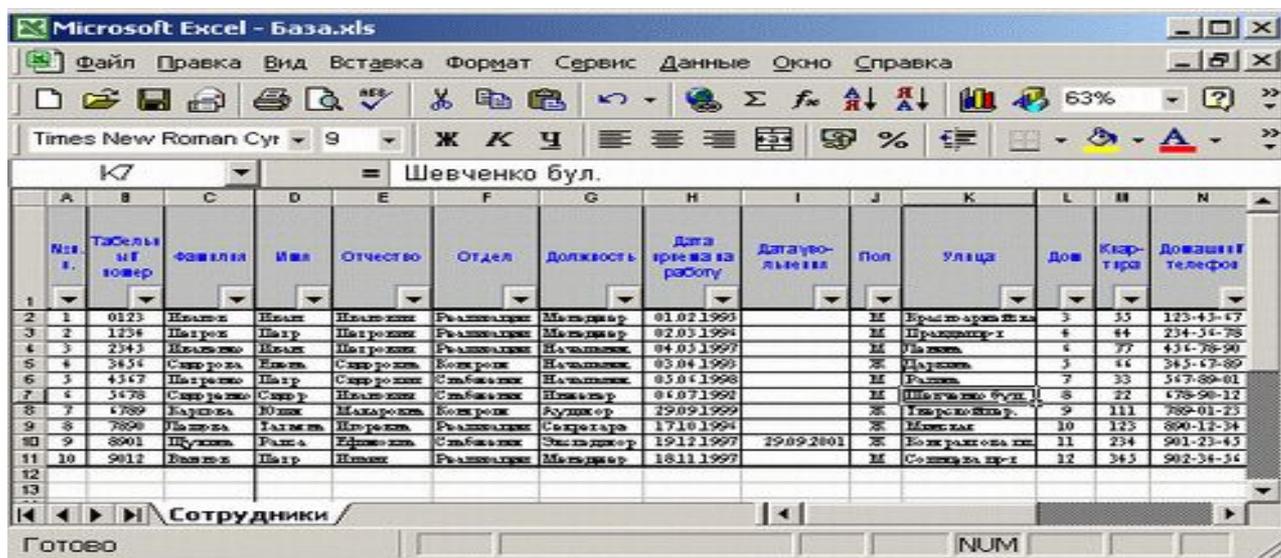
Иванов Иван Иванович работает в отделе продаж и получает заработную плату 10000р

Базовые элементы таблицы

Запись – каждая строка таблицы.

Поле – каждый столбец

Ключ – поле, которое однозначно определяет соответствующую запись.



Microsoft Excel - База.xls

Файл Правка Вид Вставка Формат Сервис Данные Окно Справка

Times New Roman Cyr 9 Ж К Ч

Шевченко бул.

№ п.п.	Табельный номер	Фамилия	Имя	Отчество	Отдел	Должность	Дата рождения	Дата приема на работу	Пол	Улица	Дом	Квартал	Домашний телефон
1	0123	Иванов	Иван	Иванович	Ремонтный	Мастер	01.02.1993		М	Бродячий перекр.	3	33	123-43-47
2	1234	Петров	Петр	Петрович	Ремонтный	Мастер	02.03.1994		М	Школьная ст.	4	44	234-54-58
3	2345	Иванов	Иван	Иванович	Ремонтный	Начальник	04.05.1997		М	Улица	4	77	434-78-90
4	3454	Сидорова	Елена	Сидоровна	Бухгалтер	Начальник	03.04.1998		Ж	Улица	3	44	345-67-89
5	4347	Петров	Петр	Петрович	Служба тех.	Начальник	05.04.1998		М	Рынок	7	33	547-89-01
6	5478	Сидорова	Светлана	Ивановна	Служба тех.	Начальник	04.07.1992		Ж	Школьная ст.	8	22	478-90-12
7	6789	Кузнецов	Юрий	Иванович	Бухгалтер	Аудитор	29.09.1999		Ж	Улицы	9	111	789-01-23
8	7890	Попов	Игорь	Игоревич	Ремонтный	Сотрудник	19.10.1994		Ж	Мостовая	10	125	890-12-34
9	8901	Щукин	Роман	Ефимович	Служба тех.	Экспедитор	19.12.1997	29.09.2001	Ж	Богородицкая ст.	11	234	901-23-45
10	9012	Иванов	Петр	Иванович	Ремонтный	Мастер	18.11.1997		М	Солнечная ст.	12	345	902-34-56

Сотрудники

Готово NUM

Основной элемент БД – таблица

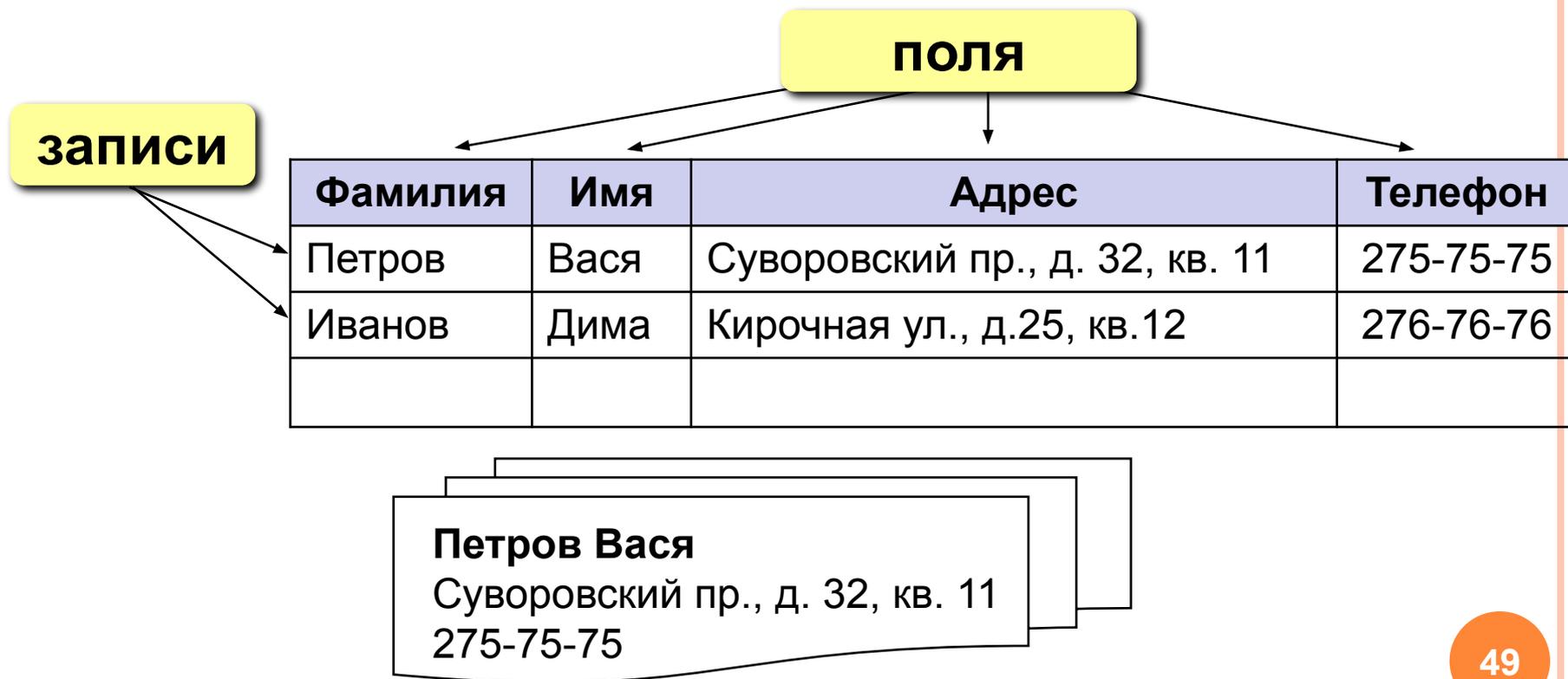
- Каждая таблица должна иметь своё *имя*.
- *Запись* – это строка таблицы.
- *Колонка* – это столбец таблицы.
- *Поле* – это атрибут строки.
- *Таблица* – информационная модель реальной системы.
- *Запись* содержит *информацию* об одном *конкретном объекте*.
- *Поле* содержит определённые *характеристики* объектов.

ЗАПИСЬ

Пол

Колонка

Пример таблицы



ТИПЫ ПОЛЕЙ

Тип данных	Использование
Текстовый	Одна строка текста (до 255 символов)
Мемо	Текст, состоящий из нескольких строк, которые затем можно будет просмотреть при помощи полос прокрутки (до 65 535 символов)
Числовой	Различные числовые данные (имеет несколько форматов: целое, длинное целое, с плавающей точкой)
Дата \ Время	Дата и время в одном из предлагаемых БД форматов
Денежный	Денежные суммы, хранящиеся с 8 знаками в десятичной части. В целой части каждые три разряда разделяются запятой.
Счетчик	Уникальное длинное целое, создаваемое БД для каждой новой записи
Логические	Логические данные, имеющие значения Истина или Ложь
Объект OLE	Картинки, диаграммы и другие объекты OLE из приложений Windows

Имя таблицы

поле

запись

Борей

Файл Правка Вид Вставка Формат Записи Сервис Окно Справка

КодСотрудник Arial 10 Ж К Ч

Сотрудники : таблица

	Код сотрудника	Фамилия	Имя	Должность	Дата рождения	Дата найма	Адрес	Город
+	1	Белова	Мария	Представитель	08-дек-1968	01-май-1992	ул. Нефтяников, 14-4	Москва
+	2	Новиков	Павел	Вице-президент	19-фев-1952	14-авг-1992	Судостроительная ул. 12-245	Москва
+	3	Бабкина	Ольга	Представитель	30-авг-1963	01-апр-1992	Крещатик, 34-55	Киев
+	4	Воронова	Дарья	Представитель	19-сен-1958	03-май-1993	ул. Пехотинцев, 1-34	Киев
+	5	Кротов	Андрей	Менеджер по продажам	04-мар-1955	17-окт-1993	Зеленый просп. 24-78	Москва
+	6	Акбаев	Иван	Представитель	02-июл-1963	17-окт-1993	Студенческая ул. 22-15	Москва
+	7	Кралев	Петр	Представитель	29-май-1960	02-января-1994	Сиреневый бульв. 11-11	Москва
+	8	Крылова	Анна	Внутренний координатор	09-января-1958	05-мар-1994	Лесная ул. 12-456	Москва
+	9	Ясенева	Инна	Представитель	02-июл-1969	15-ноя-1994	Родниковый пер. 1	Киев
*		(Счетчик)						

Главный ключ – это поле или совокупность полей, которое однозначно определяет запись в таблице

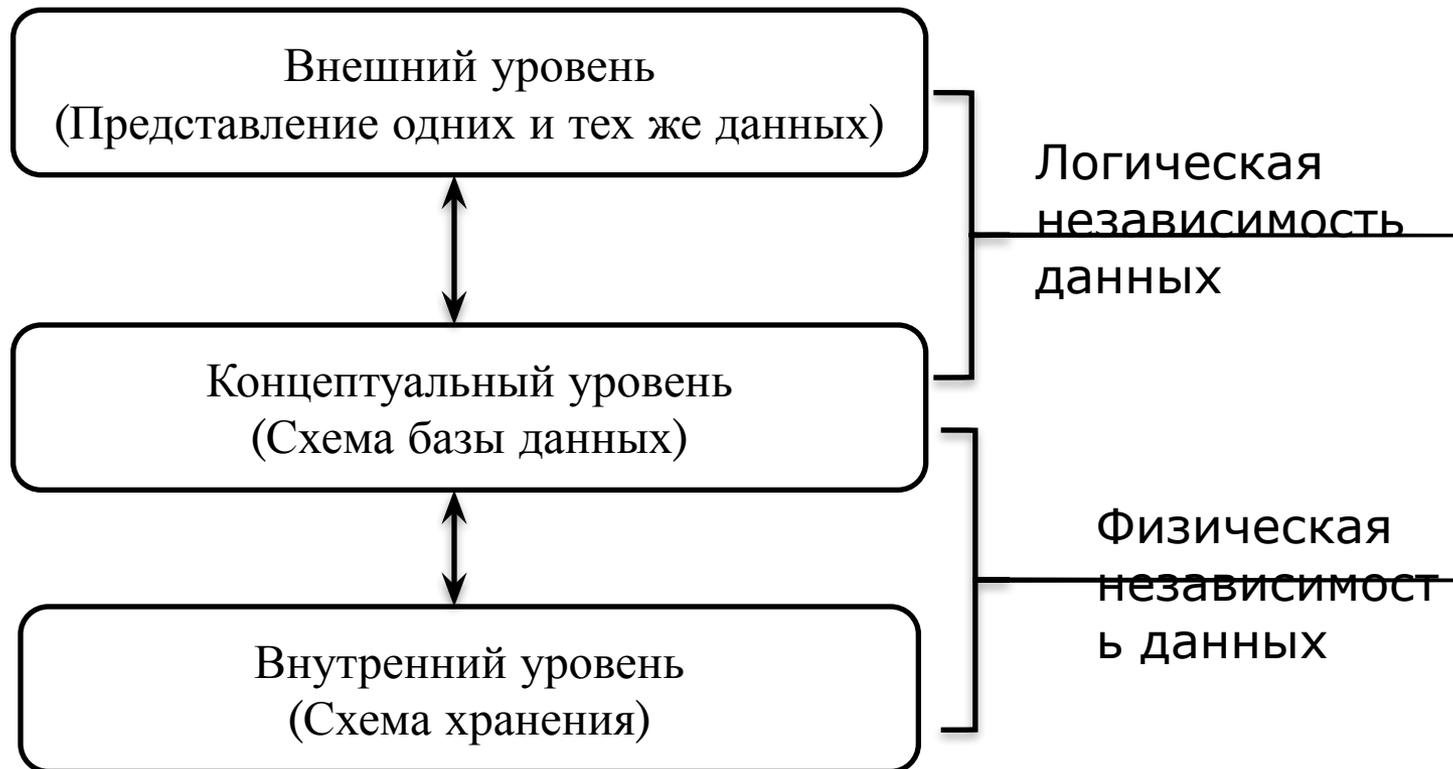
АБСТРАГИРОВАНИЕ ДАННЫХ

- Основная цель СУБД заключается в том, чтобы предложить пользователю абстрактное представление данных, скрыв конкретные особенности хранения и управления;
- Отправной точкой при проектировании базы данных должно быть абстрактное и общее описание информационных потребностей организации, которые должны найти свое отражение в создаваемой базе данных.

10.МОДЕЛЬ ПРЕДСТАВЛЕНИЯ ДАННЫХ

- Современная технология баз данных основана на концепции многоуровневой архитектуры СУБД.
- Эти идеи были сформированы в 1975 году в отчёте Американского национального института стандартов (ANSI/X3/SPARC), в котором была предложена обобщённая трёхуровневая модель архитектуры СУБД.

УРОВНИ ПРЕДСТАВЛЕНИЯ ДАННЫХ



- Внутренняя модель - представление самого низкого уровня, физическая БД. Основными компонентами являются физические блоки, хранимые записи, указатели, данные переполнения;
- Концептуальная модель - представление полного информационного содержания базы данных в абстрактной форме по сравнению со способом физического хранения данных. Это представление может полностью отличаться от представления данных отдельным пользователем;
- Внешний уровень - представление БД с точки зрения пользователей. Этот уровень описывает ту часть БД, которая относится к каждому пользователю.

КОМПОНЕНТЫ КОНЦЕПТУАЛЬНОЙ МОДЕЛИ

- элементарные данные проблемной области (или предметной области), называемые сущностями;
- элементарные данные, описывающие сущности, называемые атрибутами;
- ассоциации между экземплярами элементарных данных, называемых связями
 - 1:1 – один к одному
 - 1:N – один ко многим
 - M:N – многие ко многим

СПАСИБО ЗА ВНИМАНИЕ!