

БАЗЫ ДАННЫХ

Основные понятия



- В настоящее время жизнь человека настолько насыщена различного рода информацией, что для ее обработки требуется создание огромного количества хранилищ информации различного назначения.
- Современные информационные системы (ИС) характеризуются огромными объемами хранимых данных, сложной организацией, необходимостью удовлетворять разнообразные требования многочисленных пользователей.

Данные



Данные - это формализованное представление информации, доступное для обработки, интерпретации и обмена между людьми или в автоматическом режиме.

Информация может храниться в неструктурированном виде, например, в виде текстового документа, где данные об объектах предметной области записаны в произвольной форме:

Студент Иванов Иван родился 4 апреля 1981 года, обучается в 411 группе, номер его зачетной книжки 200205; студент Виктор Сидоров 06.08 1982 г.р., имеет зачетную книжку №200213, обучается в 413 группе

Структурированный вид



Структурированный вид хранения информации предполагает введение соглашений о способах представления данных. Это означает, что в определенном месте хранилища могут находиться данные определенного типа, формата и содержания. Указанная выше информация о гражданах Петрове, Сидорове и Иванове в структурированном виде будет выглядеть следующим образом:

№	Фамилия	Имя	Дата рождения	Учебная группа	№ зачетной книжки
1	Иванов	Иван	<i>04.04.1981</i>	411	<i>200205</i>
2	Петров	Евгений	<i>25.03.1982</i>	411	<i>200210</i>
3	Сидоров	Виктор	<i>06.08.1982</i>	411	<i>200213</i>

Представление информации в таблице - наилучший  способ структурирования данных. Все данные записаны в клеточках таблицы по определенным правилам – форматам, одинаковым для всего столбца. Все столбцы имеют названия. Кроме этого нетрудно заметить, что фамилии студентов записаны по алфавиту, при этом для записей имен и фамилий используются заглавные буквы. Каждая строка таблицы имеет порядковый номер.

Автоматизировать обработку данных, которые хранятся в неструктурированном виде сложно, а порой и просто невозможно. Поэтому вырабатывают определенные соглашения о способах представления данных. Обычно это делает разработчик базы данных. В результате все реквизиты имеют одинаковый вид и тип данных, что делает их структурированными и позволяет создать базу данных.

База данных



В широком смысле слова **база данных** - это совокупность сведений о конкретных объектах реального мира в какой-либо предметной области.

Кроме того, **база данных** – это хранилище данных для совместного использования.

При автоматизации деятельности человека происходит перенос реального мира в электронный формат.

Для этого выделяется какая-то часть этого мира и анализируется на предмет возможности автоматизации.

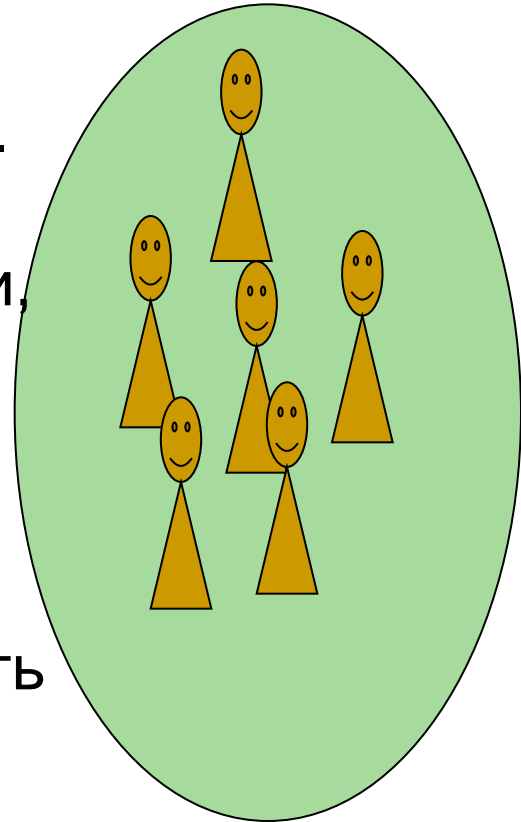
Она называется предметной областью и строго очерчивает круг объектов, которые изучаются, измеряются, оцениваются и т.д.

Предметная область

Предметная область - это часть реального мира, подлежащего изучению для организации управления и автоматизации. (Например, образование)

Объект - это элемент предметной области, информацию о котором мы сохраняем (например, студенты группы)

Реквизит (атрибут) – поименованная характеристика объекта. Он показывает, какая информация об объекте должна быть собрана (Например, Фамилия, имя, и т.д.)



База данных –
это поименованная
совокупность
структурированных
данных,
относящихся к
определенной
предметной области.



Классификация баз данных

По технологии обработки данных базы данных подразделяются на централизованные и распределенные.

Централизованная база данных хранится в памяти одной вычислительной системы, к которой подключены несколько других компьютеров.

Распределенная база данных состоит из нескольких, возможно пересекающихся или даже дублирующих друг друга частей, хранимых в различных ПК компьютерной сети. Работа с такой базой осуществляется с помощью системы управления распределенной базой данных (СУРБД).

По способу доступа к данным базы данных подразделяются на *базы данных с локальным доступом* и *базы данных с удаленным (сетевым) доступом*.

Виды моделей баз данных по логической структуре



- Реляционная (таблицы, связанные общими полями, *relatio* – отношение)
- Иерархическая – элементы БД располагаются по уровням, в порядке их подчинения
- Сетевая (подобна иерархической, но со свободными связями между элементами)



Реляционная база данных, в основу которой положена реляционная информационная система. Реляционная структура базы данных ориентирована на организацию данных в виде двумерных таблиц, называемых еще реляционными таблицами.

Каждая реляционная таблица обладает следующими свойствами:

- каждый элемент таблицы - один элемент данных;
- все столбцы в таблице однородные, т.е. все элементы в столбце имеют одинаковые характеристики и свойства;
- каждый столбец имеет уникальное имя;
- одинаковые строки в таблице отсутствуют;
- порядок следования строк и столбцов может быть произвольным.

Основные элементы реляционной модели



- **Поле** - это элементарная единица логической организации данных, которая соответствует неделимой единице информации - реквизиту.
- **Запись** - совокупность логически связанных полей, представленных одной строкой таблицы.
- **Файл (таблица)** – совокупность экземпляров записей одной структуры.

- **Ключ** (ключевое поле) - значение этого поля однозначно определяет соответствующую запись



Можно связать две реляционные таблицы, если ключ одной таблицы ввести в состав ключа другой таблицы

Так, если ключом таблицы «Книга» будет выбран «№ в каталоге», то такую таблицу можно связать, например, с таблицей «Список библиотечного фонда».

Принцип нормализации:

- в одной и той же таблице не может находиться повторяющихся полей;
- в каждой таблице ключ должен однозначно определять запись из множества записей;
- значению ключа должно соответствовать исчерпывающая информация об объекте таблицы;
- изменение значения любого не ключевого поля не должно влиять на информацию в других полях.



Между записями двух таблиц могут быть установлены следующие основные виды связей:

один к одному – 1:1 - эта связь предполагает, что в каждый момент времени одному экземпляру информационного объекта А соответствует не более одного экземпляра информационного объекта В и наоборот; например, куратор курса – курс;

один к многим – 1: N - эта связь предполагает, что одному экземпляру информационного объекта А соответствует 0, 1, 2 или более экземпляров объекта В, но каждый экземпляр объекта В связан не более чем с 1 экземпляром объекта А, например, куратор курса – студент;

многие ко многим – N:N - эта связь предполагает, что в каждый момент времени одному экземпляру информационного объекта А соответствует 0, 1, 2 или более экземпляров объекта В и наоборот, например, учебная дисциплина - студент.

Примеры моделей баз данных

Реляционная модель.



Школьный журнал 11- а класса

Таблицы журнала –

Оценки по предметам,

Сведения об учениках,

Сводная ведомость,

Сведения о занятиях на факультативах

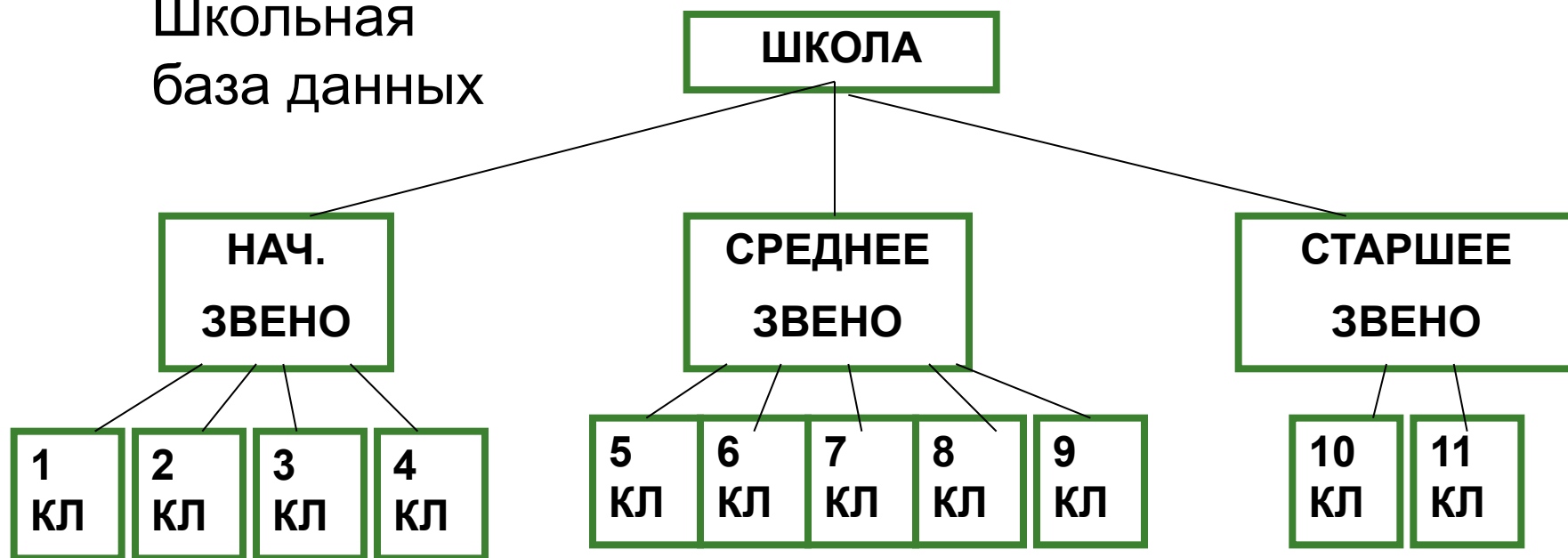
и другие

связаны ключевым полем – фамилия, имя ученика.

Иерархическая структура



Школьная
база данных





Иерархическое дерево имеет только одну вершину (корень дерева), не подчиненную никакой другой вершине и находящуюся на самом верхнем (первом) уровне.

Зависимые (подчиненные) узлы находятся на втором, третьем и т.д. уровнях.

Количество деревьев в базе данных определяется числом корневых записей.

В каждой записи базы данных существует только один (иерархический) путь от корневой записи.

Иерархическая структура в форме таблицы



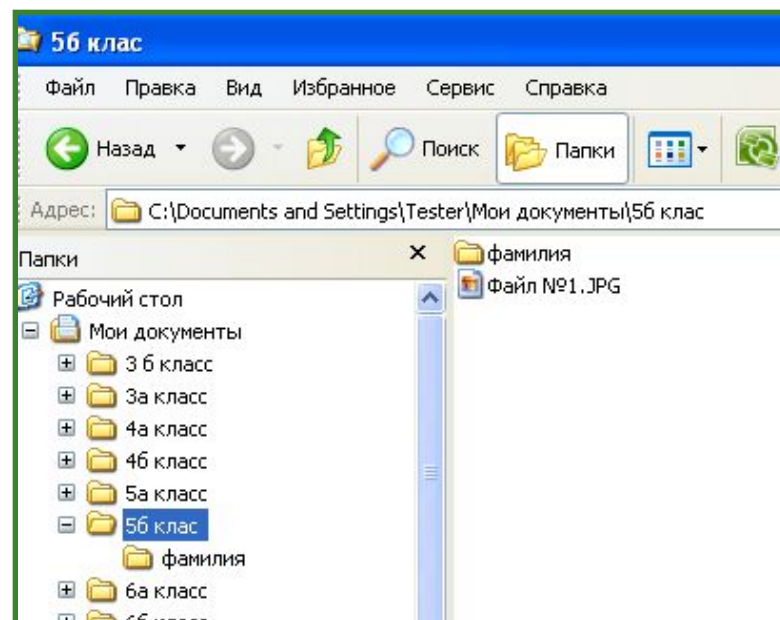
Группа биотопов	Основные биотопы	Их подразделения и объем в км²	
Свободное существование	Водная среда	Морские воды 1 370 000 000	Полносоленая морская вода Солоноватые воды Морские грунты
		Пресные воды (не считая материковых льдов) 8 300 000	Надземные воды суши 4 000 000 Подземные воды 4 000 000
	Воздушная среда	Открытые материковые воздушные биотопы 7 440 600	Лес, степь, поле, пустыня и т. д.
		Скрытые воздушные биотопы	Наземные почвы 149 000 Пещеры
Паразитическое существование	Тело других организмов		Тело животных Тело растений

Сетевая модель



Сетевая (полносвязная) база данных. В сетевой структуре базы данных при тех же основных понятиях иерархической базы данных: узел, уровень, связь - каждый элемент может быть связан с любым другим элементом.

Пример: Операционная система Windows



Основные элементы иерархической и сетевой моделей



- Уровень – линия элементов одной степени подчинения
- Узел – информационная модель элемента на данном уровне иерархии
- Связи – линии на схеме, показывающие возможные связи между элементами

Управление базами данных



ИС – информационная система — это взаимосвязанная совокупность информационных, технических, программных, технологических и других средств, а также персонала, предназначенная для сбора, обработки, хранения и выдачи информации и принятия управленческих решений.

ИПС - информационно-поисковая система - это система, обеспечивающая поиск и отбор необходимых данных в специальной базе с описаниями источников информации (индексе) на основе информационно-поискового языка и соответствующих правил поиска.

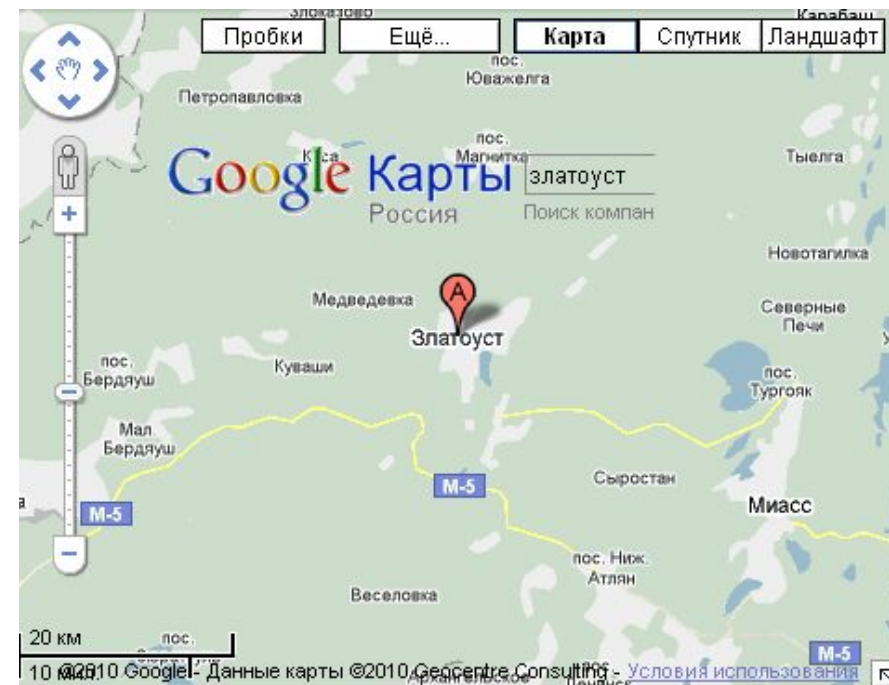
Главной задачей любой ИПС является поиск информации соответствующей информационным потребностям пользователя. Очень важно в результате проведенного поиска ничего не потерять, то есть найти все документы, относящиеся к запросу, и не найти ничего лишнего.

ГИС – геоинформационные системы



ГИС обрабатывает данные, организованные в виде геоинформационной базы данных (ГБД).

Основными компонентами ГБД являются карты, темы, таблицы, выборки, макеты, список пользователей и библиотека стилей.



	Продукт	Дата	Размер	Файл
Профессиональная ГИС				
	Профессиональная ГИС "Карта 2008" (версия 10.6.3)	26.01.2010	104 Мб	panorama10.zip
	Обновление для Профессиональной ГИС "Карта 2008" (версия 10.6.3)	11.02.2010	345 Кб	update1063.zip
	Профессиональная ГИС "Карта 2008" (версия 10.6.3, Локализованная английская версия)	22.01.2010	54,6 Мб	panorama10english.zip
	Комплекс гидрологических задач (версия 1.0, дополнительно к Профессиональной ГИС "Карта 2008")	18.12.2009	870 Кб	hydrology.zip



Система управления базами данных (СУБД) - это программный механизм, предназначенный для записи, поиска, сортировки, обработки (анализа) и печати информации, содержащейся в базе данных. В компьютерной базе данных информация представляется в виде **таблицы**, очень похожей на электронную таблицу.

Названия столбцов, представляющих «шапку» таблицы, называют **именами полей** или **реквизитами**, а сами столбцы - **полями**.

Данные в полях называют **значениями реквизитов** или **значениями полей**.

Терминология



Тип поля. Подобно электронной таблице, работающей с тремя типами полей: текстовый, числовой и формула, в таблицах используется несколько большее количество типов полей.

Длина поля – максимально возможное количество символов.

Точность (для числовых типов полей) – количество знаков после запятой.

Маска ввода – форма средства автоматизации ввода, в которой вводятся данные в поле. Например, одно и то же значение имеют поля даты: 03.03.95 или 03.03.1995, или 03-март-1995, но отличаются по формату.

Сообщение об ошибке – текстовое сообщение, которое выдается в поле при попытке ввода ошибочных данных.

Условие на значение – ограничение, используемое для проверки правильности ввода данных.

Пустое и обязательное поле – свойство поля, определяющее обязательность заполнения поля при наполнении базы данных.

Индексированное поле – дополнительное имя поля, позволяющее ускорить операции поиска и сортировки записей.

Типы полей



ТЕКСТОВЫЙ – короткий текст (до 255 символов)

МЕМО – текст из нескольких строк (до 65536 символов)

ЧИСЛОВОЙ – число любого типа

ДЕНЕЖНЫЙ – число и название денежной единицы

*ДАТА/ВРЕМЯ - дата или в заданном формате,
например 01.10.2009*

*СЧЕТЧИК – значение поля добавляется
автоматически с вводом каждой записи*

*ЛОГИЧЕСКИЙ – содержит одно из значений TRUE
(истина) или FALSE(ложь)*

*OLE – таблицы, документы, графические или
звуковые файлы*

Вопросы для повторения

- 1) Что означает словосочетание «структурирование данных» ?
- 2) Дать определение понятия «база данных» ?
- 3) Виды моделей БД ?
- 4) Основные объекты и параметры реляционной модели базы данных: запись, поле, ключ ?
- 5) Основные типы полей - числовой, дата, текстовый, MEMO, OLE, логический, денежный и другие. ?
- 6) Определение, основные функции СУБД ?
- 7) Основной признак классификации СУБД ?
- 8) Виды связей между записями двух таблиц: 1:1, 1: N, N:N ?
- 9) Что такое «маска» поиска данных? ?
- 10) Основные объекты иерархической и сетевой моделей: узел, связи, уровни. ?
- 11) Возможен ли переход из одного вида модели БД к другому? ?

Источники

- Википедия. Базы данных.
http://ru.wikipedia.org/wiki/%D0%91%D0%B0%D0%B7%D0%B0_%D0%B4%D0%B0%D0%BD%D0%BD%D1%8B%D1%85
- Информатика и ИКТ. 11 класс. Базовый уровень/ Под ред. проф. Н.В.Макаровой. — СПб: Лидер, 2010