

КУРС «БАЗЫ ДАННЫХ»



Основная литература

- Малыгина М.П. Базы данных: основы, проектирование, использование. – СПб.: БХВ-Петербург, 2004. – 512 с.
- Хомоненко А.Д., Цыганков В.М., Мальцев М. Г. Базы данных: Уч. для высших уч. заведений – СПб.: КОРОНА принт, 2004. – 736 с.

Дополнительная литература

- Плещеев В.В. Базы данных Visual FoxPro, Access, SQL SERVER, Oracle с примерами и упражнениями: Учебное пособие. – Екатеринбург, 2002.
- Харитонов И. Программирование в Access 2002. Учебный курс. – СПб.: Питер, 2002.

Информационные сетевые ресурсы

INTUIT.ru: Учебный курс – Введение в реляционные базы данных [Электронный ресурс]. Режим доступа: <http://www.intuit.ru/>

INTUIT.ru: Учебный курс – Статистические методы анализа данных [Электронный ресурс]. Режим доступа: <http://www.intuit.ru/>

INTUIT.ru: Учебный курс – Распределенные базы и хранилища данных [Электронный ресурс]. Режим доступа: <http://www.intuit.ru/>

INTUIT.ru: Учебный курс – Data Mining [Электронный ресурс]. Режим доступа: <http://www.intuit.ru/>

INTUIT.ru: Учебный курс – Проектирование хранилищ данных для приложений систем деловой осведомленности (Business Intelligence Systems) [Электронный ресурс]. Режим доступа: <http://www.intuit.ru/>

INTUIT.ru: Учебный курс – Основы проектирования приложений баз данных [Электронный ресурс]. Режим доступа: <http://www.intuit.ru/>

INTUIT.ru: Учебный курс – Базы данных [Электронный ресурс]. Режим доступа: <http://www.intuit.ru/>

INTUIT.ru: Учебный курс – Основы SQL [Электронный ресурс]. Режим доступа: <http://www.intuit.ru/>

Библиотека полнотекстовых учебников и учебных пособий по гуманитарно-экономическим и техническим дисциплинам [Электронный ресурс]. Режим доступа: <http://window.edu.ru/window/library>. – Загл. с экрана.

Историческая справка

Первые коммерческие компьютерные системы использовались в основном для ведения бухгалтерии. Они назывались **системами обработки данных**.

Файл на магнитной ленте содержал последовательность записей одинаковой структуры.

Программы обработки содержали всю информацию о структуре записей и их расположении на ленте, выполняли обмены между оперативной и внешней памятью.

Запись как единица информации:

12345	Шкаф	15 000 р.	Фирма П
-------	------	-----------	---------

Проблемы первых систем

1. Последовательный доступ к записям и, как следствие, низкая скорость их считывания и обработки, необходимость дублирования.
2. Зависимость программ от данных. Каждая прикладная программа содержала описание структур записей, сама определяла их расположение на магнитной ленте, выполняла обмены между оперативной и внешней памятью.
3. Дублирование полей в записях.

12345	Шкаф	15 000 р	Фирма П
-------	------	----------	---------

21	Фирма П	Адрес	344-344-344
----	---------	-------	-------------

Пути решения проблем

Появление магнитных дисков и прямого доступа к участку памяти.

Реализация индексно-последовательного доступа, позволяющего прямой доступ к записи по ключу.

Появление программ общего назначения, в том числе для управления базами данных.

Индексно-последовательный доступ

Индекс – служебная структура, создаваемая для быстрого поиска по значениям ключа. Ключом может быть одно или несколько полей, составное значение которых идентифицирует запись.

Индекс включает значения ключа и сведения о месте нахождения записи с этим значением ключа.

Данные		Индекс	
Название	Цена	Значение поля Название	Номер записи
Стол	100	Лампа	2
Лампа	200	Полка	5
Стул	178	Стол	1
Шкаф	322	Стул	3
Полка	320	Шкаф	4

Записи в индексе всегда упорядочены по значениям ключа.

При поиске сначала просматривается индекс для определения номера записи, а затем по номеру выбирается запись.

Для быстрого поиска в индексе используются различные методы быстрого поиска в массиве упорядоченных данных.

ОСНОВНЫЕ ПОНЯТИЯ И ОПРЕДЕЛЕНИЯ



Понятие предметной области

Предметная область – это часть реальной действительности, рассматриваемая с целью автоматизации деятельности предприятия или организации.

Предметная область представляется множеством объектов, обладающих множеством различных свойств и находящимся в бесконечном количестве взаимосвязей.

Объекты в теории базы данных принято называть сущностями, каждая из которых представляет собой набор определенных свойств – полей.

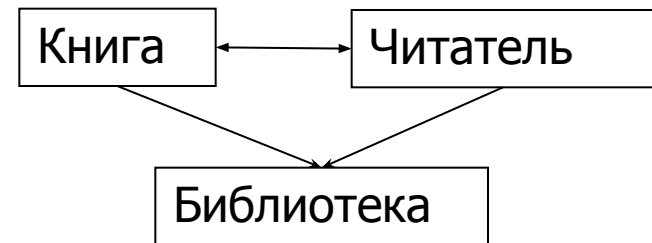
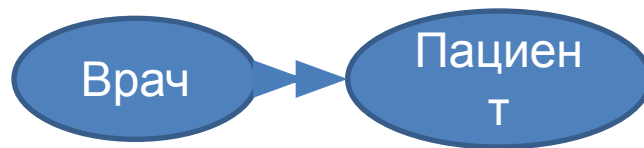
Например, сущность книга обладает такими свойствами, то есть полями как номер, автор, название, жанр, издательство, год издания и т.д.

Моделирование данных

Анализ предметной области предполагает моделирование данных, то есть выделение сущностей, связей между ними, а затем детализацию каждого из полученных объектов.

При моделировании данных используются различные методики, системы обозначений, есть стандарты.

Как правило, сущности изображаются либо в овалах, либо в прямоугольниках или ромбах, в зависимости от используемого подхода и стандарта.



При детализации объектов, всегда выделяются их свойства (характеристики), которые в последствие становятся полями сущности. Среди характеристик объекта (полей сущности) всегда выделяется *идентификатор – ключ*, по которому можно однозначно распознать запись в базе данных.

Шифр

Название

Автор

Издательство

Год издания

Отношения между объектами описываются как связи. При описании указывается смысл отношений.

Например:

- продавец принимает заказ.



Понятие базы данных

Данные – это информация, представленная в дискретном фиксированном виде, позволяющая автоматизировать ее сбор, хранение, дальнейшую обработку и передачу по каналам связи.

База данных – собрание данных, которые организованы и хранятся в соответствии с некоторым назначением [Р. Ягер].

База данных – это совокупность специальным образом организованных данных, хранимых в памяти компьютера и отражающих состояние объекта и их состояние в рассматриваемой предметной области [А. Д. Хомоненко].

База данных – совместно используемый набор логически связанных данных (и описание этих данных), предназначенный для удовлетворения информационных потребностей организации [Т. Коннолли].

База данных – это единое, большое хранилище данных, которое однократно определяется, а затем используется одновременно многими пользователями из разных подразделений. Причем база данных хранит не только рабочие данные этой организации, но и их описания. По этой причине базу данных еще называют набором интегрированных записей с самоописанием.

Отличительные признаки компьютерной базы данных:

- В базе данных хранятся не только данные пользователя, но и их описание (метаданные).
- Данные, хранящиеся в базе данных, логически структурированы. Схема базы данных содержит информацию о типах объектов предметной области, атрибутах, ограничениях и взаимосвязях. Схема также хранится в базе данных.
- База данных создается и поддерживается средствами СУБД.

Что хранится в базе данных

- данные пользователей,
- метаданные,
- служебные структуры (индексы, связанные списки и пр.),
- метаданные приложений (для настольных баз данных).

К метаданным относится:

- описание структур,
- информация о пользователях базы данных,
- форматы отображения,
- статистика обращений.

Базы данных



По архитектуре:

Файл-сервер. Предполагает выделение одной машины в сети в качестве центральной (сервер файлов), на ней хранится централизованная БД, которая используется совместно.

Клиент-сервер. Предполагается выделение сервера БД, который кроме хранения осуществляет обработку данных. Систему БД можно рассматривать как систему, состоящую из двух частей: сервер и набор клиентов.

Свойства баз данных

Восстанавливаемость. Данное свойство предполагает возможность восстановления БД после сбоя системы или отдельных видов порчи системы.

Безопасность. Безопасность БД предполагает защиту данных от преднамеренного и непреднамеренного доступа, модификации или разрушения. Применяется запрещение несанкционированного доступа, защита от копирования и криптографическая защита.

Целостность. В каждый момент времени существования БД сведения, содержащиеся в ней, должны быть непротиворечивы. Желательно отслеживать диапазон допустимых значений, соотношения между значениями в полях, особенности написания формата.

Эффективность. Свойство эффективности обычно понимается как минимальное время реакции на запрос пользователя, минимальные потребности в памяти, сочетание этих параметров.

Предельные размеры и эксплуатационные ограничения. Предельные размеры, а также другие ограничения, накладываемые эксплуатацией данной БД, могут существенно повлиять на проектное решение.

Экономичность. Затраты на обработку информации в базе данных должны быть меньше экономического выигрыша при использовании этой информации.

Переносимость. Возможность использования базы данных на разных компьютерах с различным системным программным обеспечением.

Пользователи базы данных

- **Администраторы** – люди ответственные за работу информационной системы, использующей базу данных и прикладного программного обеспечения. На стадии разработки базы данных администраторы выполняют анализ предметной области, определяют потребности пользователей, структуру будущей базы данных, средства защиты данных. Во время эксплуатации базы данных администраторы обеспечивают работу конечных пользователей.
- **Пользователи (конечные пользователи)** – это люди, которые работают с базой данных, но не причастны к сопровождению. От них не требуется специальных знаний в области вычислительной техники. Для конечных пользователей, как правило, разработаны специальные формы ввода или предоставления нужных данных.
- **Разработчики.** Выполняют проектирование и создание базы данных, а также, при необходимости выполняют её реорганизацию. Администраторы координируют работу разработчиков.

Понятие СУБД

Система управления базами данных (СУБД) – специализированная программа или комплекс программ, предназначенных для создания, ведения и совместного использования базы данных.

Предоставляемые возможности

СУБД обеспечивают постоянное хранение большого объема данных с возможностью доступа и обновления.

Пользователи и прикладные программы получают контролируемый доступ к данным посредством языка запросов.

Функции СУБД

Управление данными во внешней памяти

Создание и поддержание необходимых структур внешней памяти.

Управление буферами оперативной памяти

При работе с базой данных для увеличения скорости выполнения запросов большое количество данных находится в оперативной памяти.

Управление транзакциями

Выполняется всё или ничего.

Восстановление базы данных после сбоев

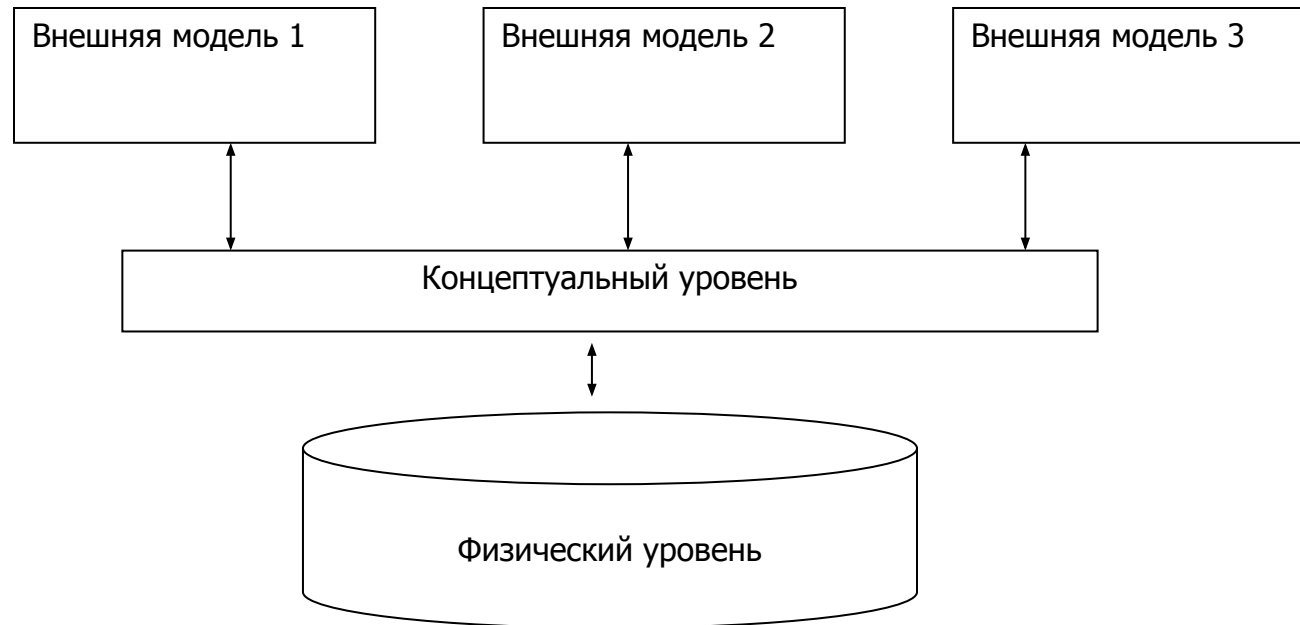
Поддержка языков

Стандартным языком наиболее распространенных в настоящее время реляционных СУБД является язык SQL (Structured Query Language).

Обеспечение безопасности

Шифрование, пароль, поддержка уровней доступа к базе данных и отдельным её элементам.

Трехуровневая система организации данных в СУБД, предложенная ANSI



Физический уровень – это собственно данные

Концептуальный уровень – представление данных, отражающее обобщенную модель предметной области.

Уровень внешних моделей – представление данных для конечного пользователя

Классификация СУБД

Системы управления базами данных значительно различаются по своим характеристикам и функциям. Их классифицируют:

- по используемой модели данных,
- по характеру использования,
- по степени распределенности.

Характер использования

Персональные (настольные) СУБД предназначены в основном для создания информационных систем, рассчитанных на одного пользователя. Они обычно включают развитые средства для создания приложений. Пример - СУБД MS Access

Многопользовательские системы позволяют организовать одновременную работу различных пользователей сети с одной базой данных. На основе таких СУБД создаются информационные системы рабочих групп и организаций.

Многопользовательские системы включают в себя *серверную* и клиентскую части. Сервер принимает запросы пользователей, выполняет их и возвращает результаты по сети.

Примеры многопользовательских СУБД – Oracle, DB2, MS SQL Server, Informix.

Компоненты системы баз данных



Расширение: * .accdb (раньше – * .mdb),
все в одном файле

Состав:

- таблицы
- формы – диалоговые окна для ввода и редактирования данных
- запросы – обращения к базе данных для выбора нужной информации или изменения базы
- отчеты – документы для вывода на печать
- макросы – средства автоматизации работы
- модули – дополнительные процедура на языке *Visual Basic*

Режимы работы: просмотр, конструктор, *SQL*



ФАЙЛ ГЛАВНАЯ СОЗДАНИЕ ВНЕШНИЕ ДАННЫЕ РАБОТА С БАЗАМИ ДАННЫХ

Режим
Вставить
Вырезать
Копировать
Формат по образцу

Режимы

Буфер обмена

Фильтр
По возрастанию
По убыванию
Удалить сортировку
Выделение
Дополнительно
Фильтр

Сортировка и фильтр

Обновить все
Создать
Сохранить
Удалить
Итоги
Орфография
Дополнительно

Записи

Все объекты A... <>

Переход в категорию

Другие

Тип объекта

Таблицы и связанные представления

Дата создания

Дата изменения

Фильтр по группам

Таблицы

Запросы

Формы

Отчеты

Все объекты Access