r

Лекция 6. Базы данных

- 6.1. История возникновения баз данных
- 6.2. Основные понятия баз данных, виды моделей и структурирование данных.
- 6.3. Модели данных, особенности их организации и поиска в них данных.
- 6.4. Структура СУБД.
- 6.5. Базы данных с графической информацией.
- (Примеры студ. работ Access)
- 6.6. Процесс создания базы данных.

6.1. История возникновения баз данных

- В широком аспекте понятие истории баз данных обобщается до истории любых средств, с помощью которых человечество хранило и обрабатывало данные. В таком контексте упоминаются, например, средства учёта царской казны и налогов в древнем Шумере (4000 г. до н. э.), узелковая письменность инков кипу, клинописи, содержащие документы Ассирийского царства и т.п. Недостатком этого подхода является размывание понятия «база данных» и фактическое его слияние с понятиями «архив» и даже «письменность».
- История баз данных в узком аспекте рассматривает базы данных в традиционном (современном) понимании. Эта история начинается с 1955 года, когда появилось программируемое оборудование обработки записей. Программное обеспечение этого времени поддерживало модель обработки записей на основе файлов. Для хранения данных использовались перфокарты.

- Оперативные сетевые базы данных появились в середине 1960-х.
 Операции над оперативными базами данных обрабатывались в интерактивном режиме с помощью терминалов. Простые индексно-последовательные организации записей быстро развились к более мощной модели записей, ориентированной на наборы. За руководство работой Data Base Task Group (DBTG), разработавшей стандартный язык описания данных и манипулирования данных, Чарльз Бахман получил Тьюринговскую премию.
- В это же время в сообществе баз данных СОВОL была проработана концепция схем баз данных и концепция независимости данных.
- Следующий важный этап связан с появлением в начале 1970-х реляционной модели данных, благодаря работам Эдгара Ф. Кодда. Работы Кодда открыли путь к тесной связи прикладной технологии баз данных с математикой и логикой. За свой вклад в теорию и практику Эдгар Ф. Кодд также получил премию Тьюринга.
- Сам термин *база данных* (англ. *database*) появился в начале 1960-х годов, и был введён в употребление на симпозиумах, организованных фирмой SDC (System Development Corporation) в 1964 и 1965 годах, хотя понимался сначала в довольно узком смысле, в контексте систем искусственного интеллекта. В широкое употребление в современном понимании термин вошёл лишь в 1970-е годы.

6.2. Основные понятия баз данных, виды моделей и структур данных

- Информационная система это совокупность программно-аппаратных средств, способов и людей, которые обеспечивают сбор, хранение, обработку и выдачу информации для решения поставленных задач.
- На ранних стадиях использования информационных систем применялась файловая модель обработки. В дальнейшем в информационных системах стали применяться *базы данных*.
- Базы данных являются современной формой организации, хранения и доступа к информации. Примерами крупных информационных систем являются банковские системы, системы заказов железнодорожных билетов и т.д.
- База данных это интегрированная совокупность структурированных и взаимосвязанных данных, организованная по определенным правилам, которые предусматривают общие принципы описания, хранения и обработки данных.

Обычно база данных создается для предметной области.

M

■ База да́нных — представленная в объективной форме совокупность самостоятельных материалов (статей, расчетов, нормативных актов, судебных решений и иных подобных материалов), систематизированных таким образом, чтобы эти материалы могли быть найдены и обработаны с помощью ЭВМ.

В определениях наиболее часто (явно или неявно) присутствуют следующие отличительные признаки:

- БД хранится и обрабатывается в вычислительной системе.
 Таким образом, любые внекомпьютерные хранилища информации (архивы, библиотеки, картотеки и т. п.) базами данных не являются.
- Данные в БД логически структурированы (систематизированы) с целью обеспечения возможности их эффективного поиска и обработки в вычислительной системе.
 Структурированность подразумевает явное выделение составных частей (элементов), связей между ними, а также типизацию элементов и связей, при которой с типом элемента (связи) соотносится определённая семантики и допустимые операции.
- БД включает схему, или метаданные, описывающие логическую структуру БД в формальном виде (в соответствии с некоторой метамоделью).

- ٧
 - База данных (БД)— организованная структура, предназначенная для хранения информации. Современные БД позволяют размещать в своих структурах не только данные, но и методы (т.е. программный код), с помощью которых происходит взаимодействие с потребителем или другими программно-аппаратными комплексами.
 - Системы управления базами данных (СУБД) комплекс программных средств, предназначенных для создания структуры новой базы, наполнения ее содержанием, редактирования содержимого и визуализации информации.
 - Под визуализацией информации базы понимается отбор отображаемых данных в соответствии с заданным критерием, их упорядочение, оформление и последующая выдача на устройство вывода или передача по каналам связи.



Виды моделей

Наборы принципов, которые определяют организацию логической структуры хранения данных в базе, называются моделями данных.

Существуют 4 основные модели данных:

- списки (плоские таблицы)
- реляционные базы данных
- иерархические структуры
- сетевые структуры



- В течение многих лет преимущественно использовались плоские таблицы (плоские БД) типа списков в Excel.
- В настоящее время наибольшее распространение при разработке БД получили реляционные модели данных.
- Реляционная модель данных является совокупностью простейших двумерных таблиц – отношений (англ. relation).

Простейшая двумерная таблица определяется как отношение (множество однотипных записей, объединенных одной темой).

От термина *relation* (отношение) происходит название реляционной модели данных.

- В **реляционных** БД используется несколько двумерных таблиц, в которых строки называются записями, а столбцы полями, между **записями** которых устанавливаются **связи**.
- Этот способ организации данных позволяет данные (записи) в одной таблице связывать с данными (записями) в других таблицах через уникальные идентификаторы (ключи) или ключевые поля.

M

Структурирование данных

Структура данных (в информационном смысле) — это представление пользователя о данных, не зависящее от способа их хранения. Структура данных характеризует типы данных и правила их взаимосвязи (отношений).

Тип данных образует множество значений, которые могут принимать соответствующие ему данные.

Отношение — одно из основных понятий современной информатики. Бывают: аналитические, ассоциативные, парадигматические, синтагматические, подчинения.

Парадигматическое отношение представляет собой семантическое (смысловое) отношение, существующее между словами естественного или информационного языка.

٧

Парадигматическое отношение представляет собой семантическое (смысловое) отношение, существующее между словами естественного или информационного языка. Оно также связывает слова, обозначающие предметы, между которыми существует постоянная связь.

Парадигматические отношения могут быть заданы графически в виде схем, в которых отношения показаны стрелками (графовая модель).

Частным случаем парадигматического отношения являются аналитические и ассоциативные отношения.

Аналитическим называется отношение между понятиями, существующее вследствие постоянной связи между соответствующими классами предметов и вытекающее из определений сопоставляемых понятий (мансарда, чердак).

М

Ассоциативные отношения бывают двух видов: отношение подчинения - соответствует родовидовому отношению между словами, понятиями или предметами («класс — подкласс» или «целое — часть»).; причинно-следственное (пространственное, временное) — любое устойчивое отношение между предметами, которое целесообразно учитывать при информационном поиске.

Структурирование данных задают прежде всего с помощью различного рода отношений порядка (упорядоченности). *Простейший* вид упорядоченности задают обычной нумерацией данных с помощью последовательности чисел.

Например, Xi, где і пробегает целые числа от m до n, идентифицирует упорядоченный набор данных, называемый обычно *одномерным массивом*.

v

Двухиндексный идентификатор X_{ij} идентифицирует двумерный массив и т.д.

В упорядоченных таким образом массивах возникают отношения следования. Так, следующим по индексу j для элемента X_{ij} будет элемент $X_{i,j+1}$, а предыдущим — $X_{i,j-1}$. Если индекс j пробегает значения от m до n, то для m не существует предыдущего, а для n — следующего значения индекса.

Если значения индексов задаются константами, то это прямоугольный массив. Если задано соотношение m <= i <= j <= n, то массив треугольный.

Массивы, состоящие из элементов одного и того же типа, называются *однородными*. Одномерные однородные массивы называются *векторами*, двумерные - *матрицами*.

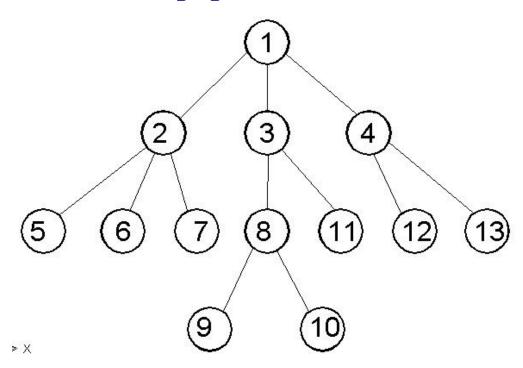
6.3. Модели данных, особенности их организации и поиска в них данных

Более сложные, составные структуры данных, представленные в виде фиксированной системы понятий и правил для описания структуры, называются *моделями данных*.

Распространены графовые и реляционные модели данных.

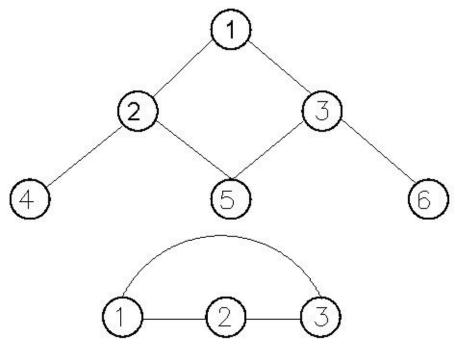
В графовых моделях структуру данных изображают и описывают в виде графа, в котором узлами (вершинами) являются типы данных, а дугами (связями) - отношения между ними. Графовый подход используют для описания иерархических и сетевых моделей данных.





Верхний узел называется *корнем*, *родовым или старшим узлом*. Узлы, из которых выходят дуги на нижний уровень, называются *исходными*, в которые входят дуги с высшего уровня — *порожденными*. Вершины нижнего уровня, не имеющие порожденных узлов, называются *листьями*. Характерно отношение 1:М (один ко многим).

Особенности сетевых моделей данных



В сетевых моделях данных порожденный узел может иметь более одного исходного, т.е. Один тип записи является членом более чем одного типа набора. Для сетей характерны отношения М:1 и М:М.



Реляционные модели данных предложены в 1970 г. Основаны на представлении данных в виде отношений, которые могут подвергаться нормализации – пошаговому процессу приведения их к двумерной табличной форме. К такой же форме могут быть приведены и иерархические и сетевые отношения. В таблице каждая строка которой соответствует значениям свойств (атрибутов), которыми обладает объект данного типа; каждый из столбцов соответствует множеству значений, которые принимает некоторый атрибут этого типа. Отношение есть множество векторов из n элементов – кортежей (X₁...X_n), где n - число столбцов, называемое степенью отношения. Совокупность значений одного атрибута (соответствующая столбцу таблицы) называется доменом.

Для описания отношений и манипуляций над ними в реляционной модели данных используется строгий математический язык, основанный на алгебре отношений (реляционной алгебре) и исчислении отношений (реляционное исчисление).

Операции реляционной алгебры позволяют вырезать отдельные домены из отношения, объединять отношения, причем, в результирующем отношении вырезаются совпадающие строки и др.

Обозначение панели	Вид	отделки	Количество бетона	Количество стали

Поиск данных в реляционной БД происходит с помощью индексных файлов. *Индексный файл*, как правило, представляет собой отдельную таблицу, содержащую ключ каждой записи и ее адрес на запоминающем устройстве.

м

6.4. Структура СУБД

В состав большинства СУБД входят три основных компонента: командный язык, интерпретирующая система или компилятор для обработки команд и интерфейс пользователя.

Командный язык служит для выполнения операций над данными, позволяет создавать прикладные программы, оформлять на экране и печатать формы ввода и вывода.

Для превращения текстовой команды в машинный код используются *интерпретаторы и компиляторы*. Первые — по очереди преобразует команды в исполнимый код перед их непосредственным выполнением, - преобразуют сначала всю программу (компилируют) в серию машинных команд и только после этого выполняют ее.

М

К числу СУБД реляционного типа относятся хорошо известные системы: Access, dBase, Clipper, FoxBASE, R:BASE, Paradox, FoxPro? Oracle и т.д. Во всех этих СУБД записи и поля имеют ограничение на число записей (128...10024) и длину (обычно 4000...5000 байт). Исключение составляет поле Мето, в котором хранят текстовую информацию большого объема.

Длина поля зависит от его типа. (Поля могут быть целыми, вещественными, строчными, логическими, типа «дата» и т. д.) В любое поле можно вписать только информацию именно этого типа.

Задание типа поля определяет множество объектов, множество операций над ними и множество правил, указывающих допустимое сочетание или применимость операций к тому или иному объекту данных.

7

Основные функции СУБД:

- определение данных (описание структуры баз данных)
- обработка данных
- управление данными

Прежде чем заносить данные в таблицы, нужно определить структуру этих таблиц. Под этим понимается не только описание наименований и типов полей, но и ряд других характеристик (например, формат, критерии проверки вводимых данных).

Кроме описания структуры таблиц, обычно задаются *связи* между таблицами. Связи в реляционных базах данных определяются по совпадению значений полей в разных таблицах:

• отношение «*один-ко-многим*» - клиенты и заказы (одной записи в таблице, содержащей сведения о клиентах, может соответствовать несколько записей в таблице заказов этих клиентов)

- м
 - отношение «*многие-ко-многим*» преподаватели и курсы лекций (преподаватель может читать несколько курсов, но и один курс может читаться несколькими преподавателями).
 - отношение «*один-к-одному»*. Как правило, это бывает в двух случаях: запись имеет большое количество полей, и тогда данные об одном типе объектов разносятся по двум связанным таблицам, или нужно определить дополнительные атрибуты для некоторого количества записей в таблице, тогда создается отдельная таблица для этих дополнительных атрибутов, которая связывается отношением "один-к-одному" с основной таблицей.



- Любая СУБД позволяет выполнять четыре простейшие операции с данными:
- добавлять в таблицу одну или несколько записей;
- удалять из таблицы одну или несколько записей;
- *обновлять* значения некоторых полей в одной или нескольких записях;
- *находить* одну или несколько записей, удовлетворяющих заданному условию.
- Под управлением данными обычно понимают защиту данных от несанкционированного доступа, поддержку многопользовательского режима работы с данными и обеспечение целостности и согласованности данных.

http://www.taurion.ru/access/1/4

6.5. Базы данных с графической информацией

Данные в БД хранятся в таблицах, связанных между собой с помощью идентификаторов. Таблицы могут содержать несколько сотен тысяч записей. Просматривать такие таблицы, особенно, если в них много ссылок друг на друга, достаточно сложно. Удобным инструментом, который позволит извлекать из таблиц и выводить на экран нужные данные, являются ЗАПРОСЫ.

Формат запроса

Стандартный запрос на выборку:

- 1. определить, какие таблицы потребуются;
- 2. какие поля в этих таблицах вас интересуют и
- 3. (если есть) условие отбора.

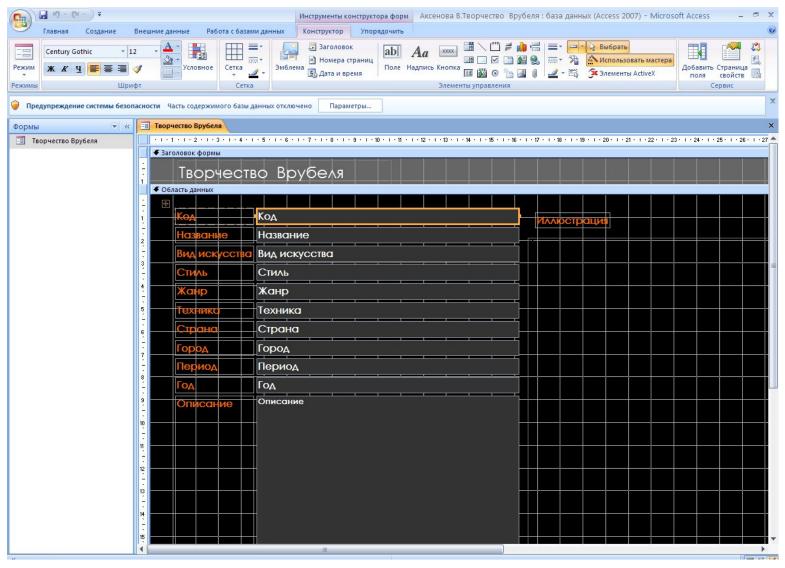
Принципы проектирования:

- 1. В каждой ячейке располагается минимальная единица информации.
- 2. Локальность хранения данных: если значения в каком-то поле регулярно повторяются, их следует выносить в отдельную таблицу, а в поле оставлять ссылку.
- 3. Объединение таблиц: если несколько сущностей имеют почти идентичные схемы данных, их следует объединить в одну таблицу.

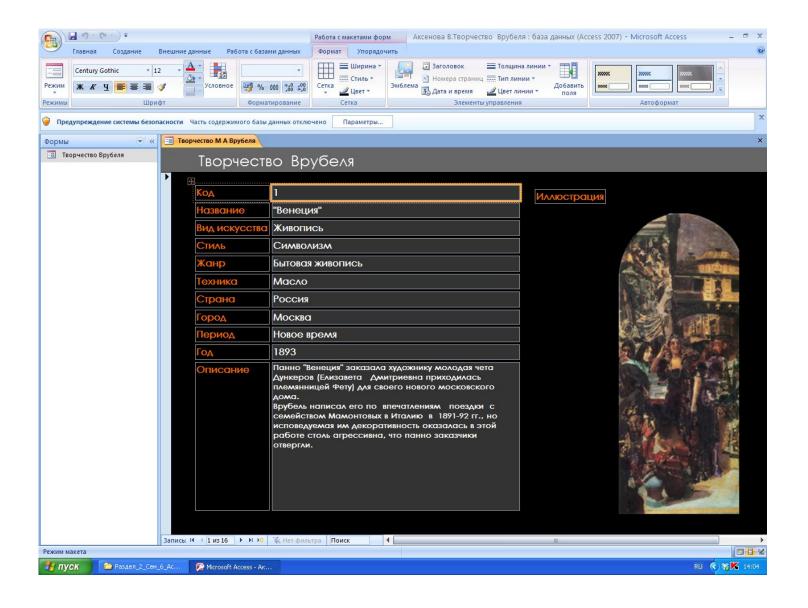
Правила формирования:

- 1. У каждой таблицы первое поле ID типа Счетчик.
- 2. Правила именования: по-английски без пробелов (каждое слово с большой буквы) + приставка tb...
- 3. Два типа связей: **один-ко-многим** (создаем полессылку), **многие-ко-многим** (создаем дополнительную таблицу).

Примеры студенческих работ – Access Режим - конструктор



Режим – макет (1)



Творчество Врубеля

Код	6		
Название	"Демон сидящий"		
Вид искусства	Живопись		
Стиль	Символизм		
Жанр	Портрет		
Техника	Масло		
Страна	Россия		
Город	Москва		
Период	Новое время		
Год	1890		
Описание	Сюжет картины навеян поэмой Лермонтова «Демон». Врубель так писал о своей работе: Демон — дух не столько злобный, сколько страдающий и скорбный, при всем этом дух властный, величавый Демон — образ силы человеческого духа, внутренней борьбы, сомнений. Трагически сцепив руки, Демон сидит с печальными, направленными вдаль огромными глазами, в окружении невиданных цветов. Фоном картины является горная местность в алом закате. Композиция подчёркивает стеснённость фигуры демона, будто бы зажатого между верхней и нижней перекладинами рамы. Картина написана в индивидуальном стиле Врубеля с эффектом кристаллических граней, что делает его картины более похожими на витражи или панно.		

Иллюстрация



Различные виды макетов – вариант 2





Произведения Альфонса Мухи



Zodiac Зодиак

Техника

Размер

Цветная литография

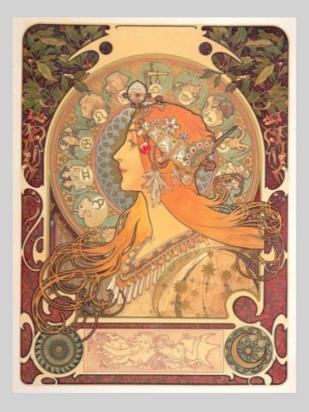
65,7 x 48,2 cm

Галерея:

Mucha Museum

Описание

Один из самых известных дизайнов Мухи, Зодиак, первоначально был выпущен Шампенуа в виде настенного календаря. Он тут же был отобран издателем La plume для журнального календаря. Знаки Зодиака окружают прелестное лицо, помещенное в центре. Впечатляет его величественность, подчеркнутая королевским нарядом и украшениями. Позже Муха на основе дизайна сделал скульптуру La Nature. Известны девять вариантов этой литографии. Один из вариантов был напечатан без текста и служил декоративным панно.



15

Новое время

Изобразительное искусство

19

1896

Европа

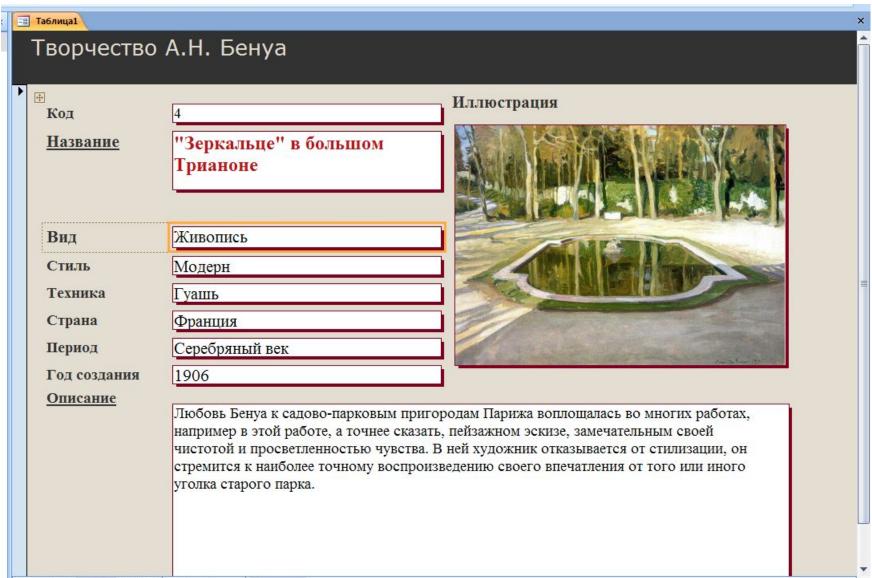
Франция

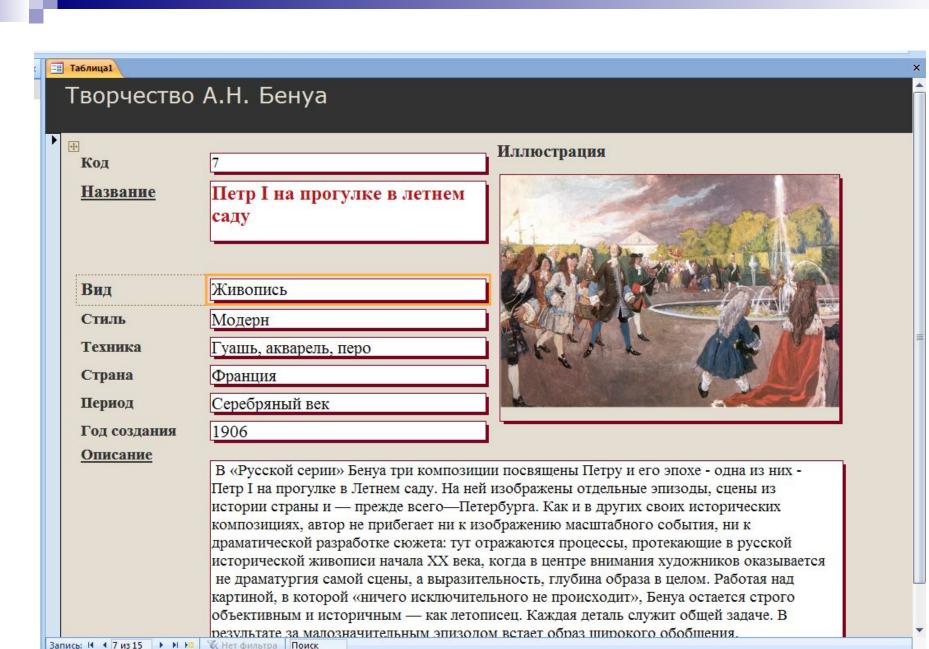
Париж

Арт Нуво (модерн)

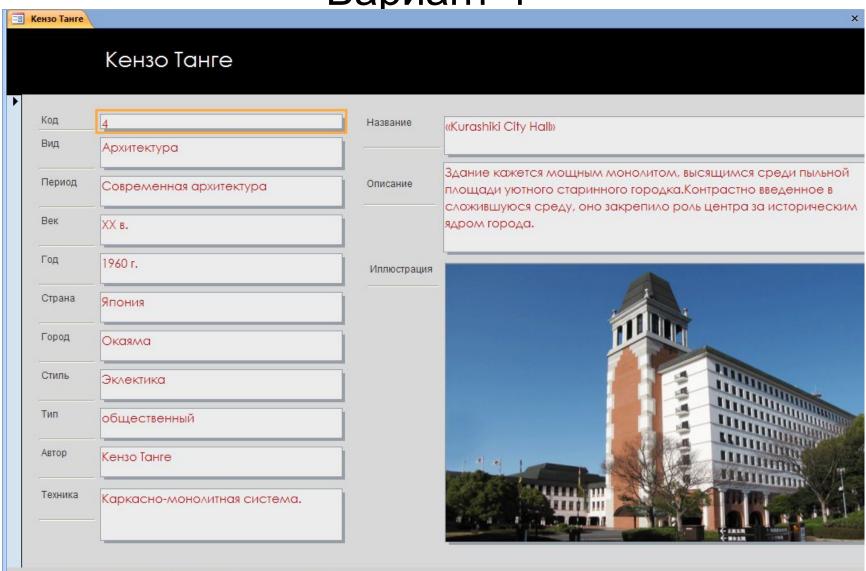
Декоративное панно

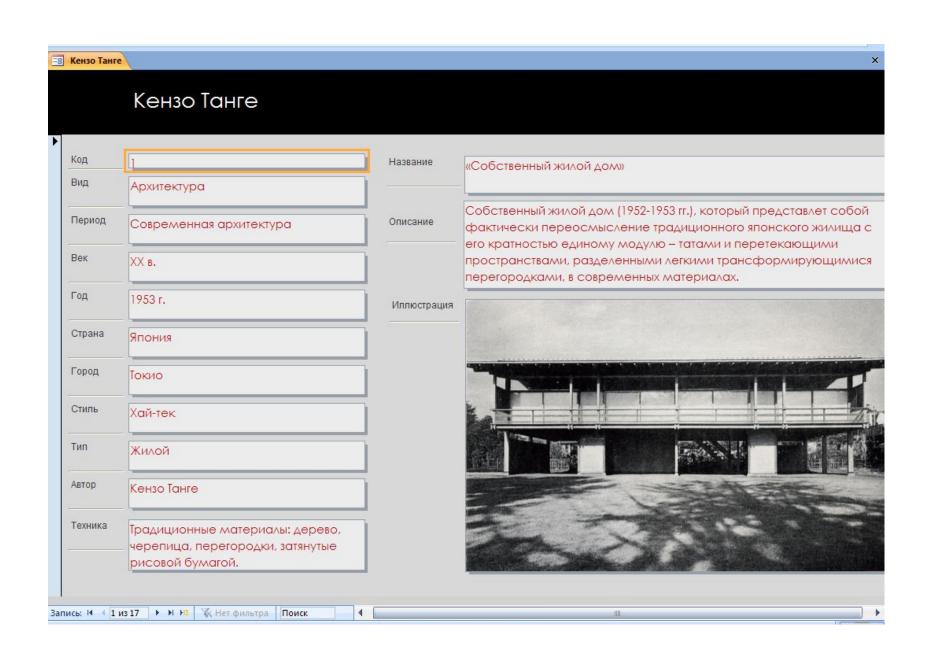
Вариант 3





Вариант 4

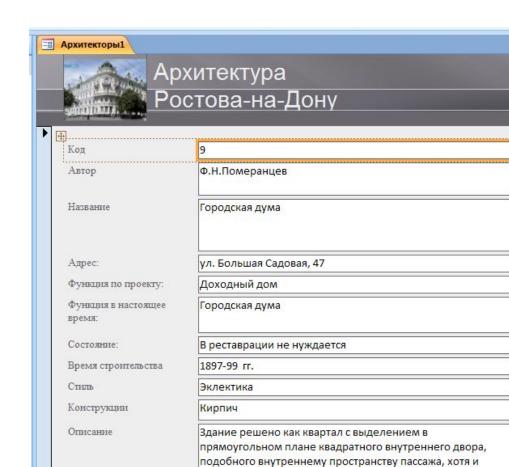






Здание расположено в средине квартала по ул. Большой Садовой между пр. Чехова и пер. Университетским. Оно Г-образное в плане, со скатной крышей, внутренним двором и въездом. Здание входит в исторический ансамбль центральной части г. Ростова-на-Дону. Является ярким представителем доходных домов начала ХХ века.

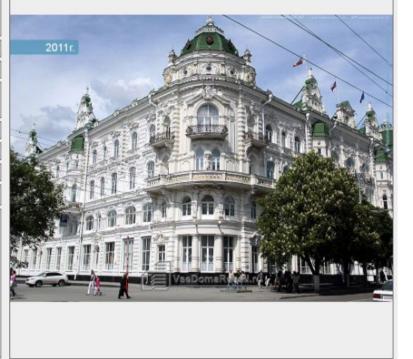
Описание



другого по функции. Над пространством двора было

запроектировано шатровое покрытие.

Иппострация





Архитектура Ростова-на-Дону

Коп

Автор

Щуко В.А., Гельфрейх В.Г.

Название

Театр им. Горького. Ростов-на-Дону

Адрес:

пл. Театральная, 1

Функция по проекту:

Театр Театр

Функция в настоящее

время:

Состояние:

В реставрации не нуждается

Время строительства

1930-35гг.

Стиль

Конструкции

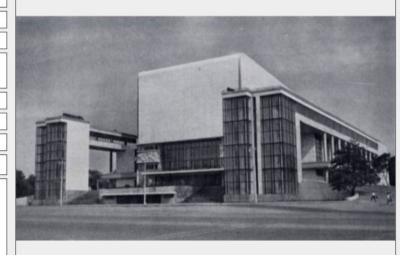
Описание

Кирпич

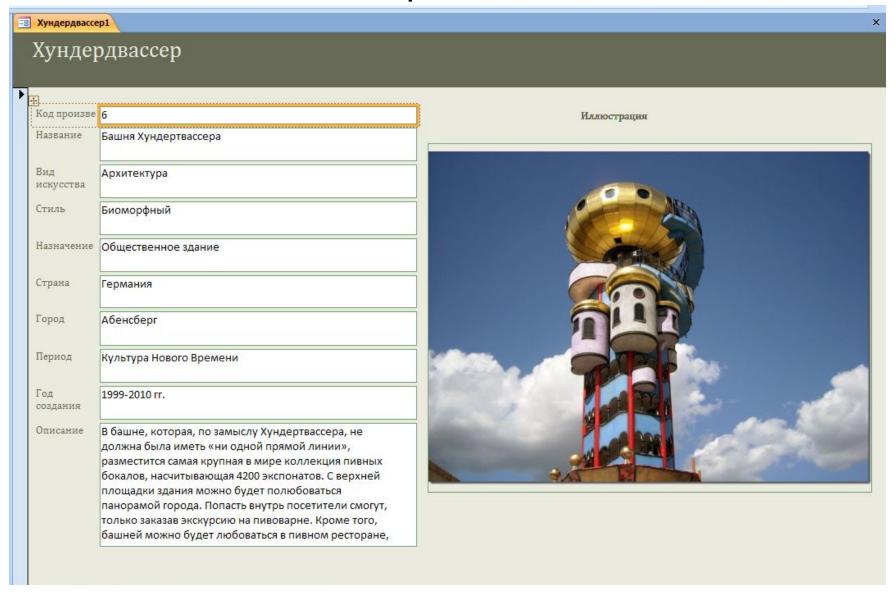
Конструктивизм

Строительству театра предшествовал Всесоюзный открытый конкурс, на котором были представлены 25 проектов, 6 из них получили премии. Первой премией был отмечен проект под девизом «Красный мак», выполненный московскими архитекторами Г.Б. и М.Г. Бархиными при участии Б.Г. Бархина. Вместе с тем к строительству был принят проект, разработанный московскими архитекторами В.А. Щуко и В.Г. Гельфрейхом, присланный после окончания срока конкурса, но привлекший внимание властей города

Иппострация



Вариант 6



٧

6.6. Процесс создания баз данных

Концепцию, в рамках которой удобно и полезно рассматривать развитие системы БД во времени, называют *жизненным циклом*.

Жизненный цикл БД делят на 2 фазы: 1) анализ и проектирование; 2) реализация и функционирование.

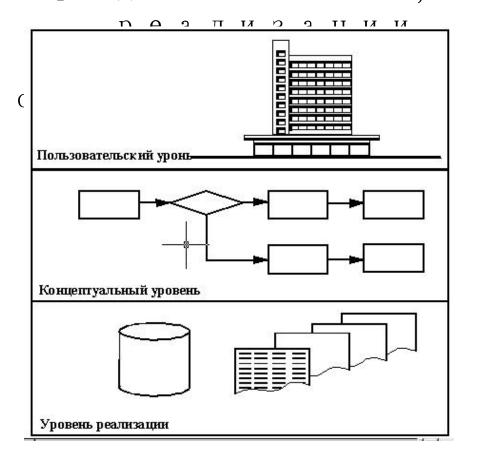
В течение первой фазы осуществляется сбор требований пользователей и проектирование БД, в течение второй – компьютерная реализация и использование БД для решения прикладных задач.

Основная цель проектирования БД – обеспечение пользователей точными данными за приемлемое время.

М

Принято рассматривать используемые для описания предметной области данные в виде трехуровневой схемы: внешнее представление,

Внешнее представление в н у т р е н н е е оанных является (физический ур р совокупностью требований конкретной прикладной функции. О н о д е л и т с я н а пользовательский и концептуальный уровни.





На концептуальном уровне определяют:

сущности (личности, факты, объекты); атрибуты (данные, описывающие сущности); связи (отношения между атрибутами).

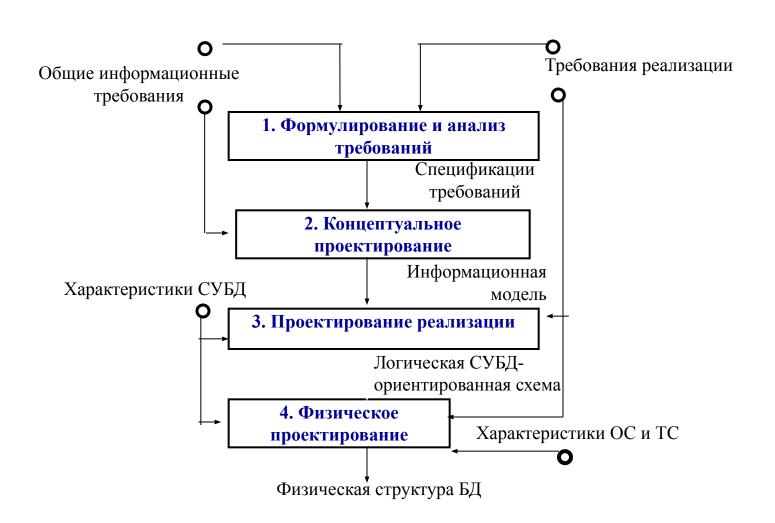
На *уровне реализации* выбирают подход к структурированию (графовый или реляционный) и модель данных, а также формируют:

записи; элементы данных; связи между записями.

На физическом уровне представления создаются:

блоки; указатели; данные переполнения; группировка данных.

Фаза анализа и проектирования БД



Фаза реализации и функционирования

Реализация БД подразумевает создание (наполнение) базы и разработку прикладных программ, а также загрузку БД. Загрузка связана с преобразованием имеющихся данных из формы логической и физической структуры в новую форму, соответствующую результатам проектирования БД.

Анализ функционирования и поддержка осуществляются для регистрации (сбора) и статистической обработки данных о функционировании системы.

Этап модернизации и адаптации предусматривает внесение в проект изменений, возникающих при появлении новых требований, полученных в результате анализа.

Реорганизация БД — это осуществление любых действий, связанных с изменением ее логической и физической структуры.