## Тема 2. Базы данных и СУБД

*Цель: Дать понятие и определить компоненты и структурные элементы базы, банка данных и системы управления базами данных. Определить основные преимущества банков данных и требования, предъявляемы к ним.*

План

1. Понятие базы данных и системы управления базой данных
2. Банки данных:

2.1 Понятие банка данных

2.2 Преимущества банка данных

2.3 Предпосылки широкого использования БнД

2.4 Требования, предъявляемые к банкам данных

1. Структура системы базы данных (компоненты БнД)
2. Функции прикладных программ
3. Системы управления базами данных (СУБД)

5.1 Функции СУБД

5.2 СУБД, как основа программных средств банка данных

5.3 Языковые средства СУБД

1. Компоненты базы данных
2. Структурирование данных
3. Структурные элементы базы данных
4. Уровни представления данных

## 1. Понятие базы данных

В широком смысле слова ***база данных*** — это совокупность сведений о конкретных объектах реального мира в какой-либо предметной области. Под предметной областью принято понимать часть реального мира, подлежащего изучению для организации управления и в конечном счете автоматизации, например, предприятие, вуз и т.д.

**2. Банки данных**

**2.1 Понятие банка данных**

***Банк данных*** (БнД) является современной формой организации хранения и доступа к информации. Существует множество определений банка данных. В «Общеотраслевых руководящих материалах по созданию банков данных» (М.: ГКНТ, 1982) дано следующее определение:

***Банк данных*** - это система специальным образом организованных данных (баз данных), программных, технических, языковых, организационно-методических средств, предназначенных для обеспечения централизованного накопления и коллективного многоцелевого использования данных».

**2.2 Преимущества БнД**:

Особенности «банковской» организации данных определяют их основные преимущества перед «небанковской» организацией.

* Наличие единого целостного отображения;
* Централизованное управление данными;
* Возможность более полной реализации принципа независимости прикладных программ от данных, чем это возможно при организации локальных файлов;
* Наличие в составе СУБД средств, ориентированных на разные категории пользователей.

**2.3 Предпосылки широкого использования БнД**

Очевидные преимущества БнД и объективные предпосылки их создания привели к широкому их использованию. К числу предпосылок применения относятся следующие:

• объекты реального мира находятся в сложной взаимосвязи между собой. Это приводит к необходимости, чтобы их информационное отражение также представляло единое взаимоувязанное целое;

• информационные потребности различных пользователей существенно пересекаются, что делает целесообразным использование единых баз данных и обеспечение доступа к ним разных пользователей;

• функции создания и ведения информационного фонда и предоставления нужных данных являются универсальными, общими при решении разнообразных задач. Создание специализированных программных средств для управления данными приводит к повышению уровня выполнения этих функций и сокращению трудоемкости создания информационных систем;

• современный уровень развития технического и программного обеспечения, а также теории и практики построения информационных систем позволяют создавать эффективные БнД.

**2.4 Требования, предъявляемые к банкам данных**

Особенности «банковской» организации данных позволяют сформулировать основные требования, предъявляемые к БнД:

• адекватность отображения предметной области (полнота, целостность и непротиворечивость данных, актуальность информации, т. е. ее соответствие состоянию объекта на данный момент времени);

• возможность взаимодействия пользователей разных категорий и в разных режимах, обеспечение высокой эффективности доступа для разных приложений;

• дружелюбность интерфейсов и малое время на освоение системы, особенно для конечных пользователей;

• обеспечение секретности и конфиденциальности для некоторой части данных; определение групп пользователей и их полномочий;

• обеспечение взаимной независимости программ и данных;

• обеспечение надежности функционирования БнД; защита данных от случайного и преднамеренного разрушения; возможность быстрого и полного восстановления данных в случае их разрушения; технологичность обработки данных, приемлемые характеристики функционирования БнД (стоимость обработки, время реакции системы на запросы, требуемые машинные ресурсы и др.).

**3. Структура системы базы данных**

На рисунке 1 показаны четыре основных элемента базы данных. Слева показано, как пользователи выполняют свои задачи с помощью базы данных. Они вводят новые данные, модифицируют существующие и удаляют ненужные. Кроме того, они разными способами читают данные: посредством форм, запросов и путем генерации отчетов.

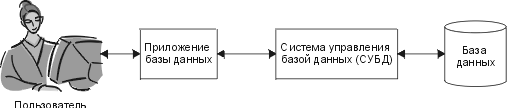


Рисунок 1 - Компоненты системы базы данных

### **4. Функции прикладных программ**

На рисунке 2 перечислены функции прикладных программ баз данных и СУБД. В первую очередь приложение создает и обрабатывает формы. Веб-приложение, например, генерирует HTML и другие конструкции веб-форм, которые заставляют форму отображаться на компьютере пользователя. Когда пользователь заполняет форму и посылает данные обратно, приложение определяет, какие таблицы данных нуждаются в модификации, и посылает запросы к СУБД, чтобы вызвать необходимую модификацию. Если во время этого процесса возникают ошибки, приложение получает сообщение об ошибке и генерирует подходящее сообщение для пользователя или осуществляет какое-нибудь другое действие.

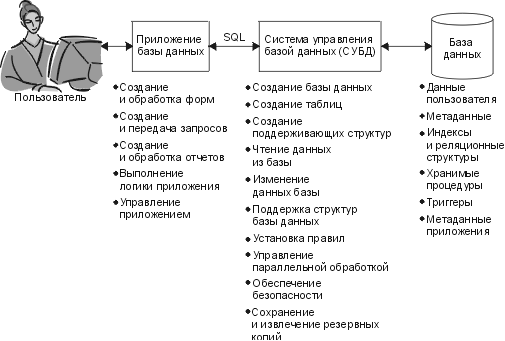


Рисунок 2 - Функции и содержимое компонентов обработки базы данных

### **5. Системы управления базами данных (СУБД)**

### **5.1 Функции СУБД**

***Система управления базами данных (СУБД)*** — это комплекс программных и языковых средств, необходимых для создания БД, поддержания их в актуальном состоянии и организации поиска в них необходимой информации.

Функции СУБД перечислены на рисунке 2. СУБД используется для создания самой базы данных, таблиц и других поддерживаемых структур - например, индексов.

**5.2 СУБД – как основа программных средств банка данных**

Основу программных средств БнД представляет СУБД. .В ней можно выделить ядро СУБД, обеспечивающее организацию ввода, обработки и хранения данных, а также другие компоненты, обеспечивающие настройку системы, средства тестирования, утилиты, обеспечивающие выполнение вспомогательных функций, таких, как восстановление баз данных, сбор статистики о функционировании БнД и др. Важной компонентой СУБД являются трансляторы или компиляторы для используемых ею языковых средств.

Подавляющее большинство СУБД работает в среде универсальных операционных систем и взаимодействует с ОС при обработке обращений к БнД. Поэтому можно считать, что ОС также входит в состав БнД.

СУБД предназначена для централизованного управления базой данных в интересах всех работающих в этой системе.

По степени универсальности различают два класса СУБД:

1) •системы общего назначения; 2) •специализированные системы.

**5.3 Языковые средства СУБД**

Языковые средства СУБД являются важнейшей компонентой банков данных, так как в конечном счете они обеспечивают интерфейс пользователей разных категорий с банком данных. Языковые средства большинства СУБД относятся к языкам четвертого поколения (к первому поколению языков относят машинные языки, ко второму — символические языки ассемблера, к третьему — алгоритмические языки типа PL, COBOL и т. п., которые в 60-е годы назывались языками высокого уровня, но уровень которых гораздо ниже, чем у языков четвертого поколения).

**Программные средства** СУБД подразделяют на:

* ***ядро СУБД****,* которое обеспечивает ввод, вывод, обработку и хранение данных в БД;
* ***трансляторы****,* обеспечивающие перевод языка СУБД на некоторый внутренний язык, используемый ядром;
* ***утилиты****,* которые служат для настройки системы, отладки программ, архивирования и восстановления БД, сбора статистики;
* ***прикладные программы****,* которые служат для обработки запросов к БД.

Операционную систему иногда включают в состав банка данных, так как СУБД тесно взаимодействует с ОС в процессе работы.

***Языковые средства*** обеспечивают взаимодействие пользователей с БД. Язык обычно включает в себя средства спецификации данных, отчетов; экранных форм, запросов и процедурные средства для описания последовательности решения задач. Язык СУБД может быть универсальным языком программирования с включением специфического подъязыка для работы с БД, например, языки универсальных систем программирования DELPHI, Visual Basic 5, Visual C++ включают язык SQL. Другие СУБД имеют специализированные языки, например, dBASE, FoxPro, Clipper, Paradox, Access. Некоторые СУБД используют только язык SQL (SQL- серверы).

### **6. Компоненты базы данных**

На рисунке 2 компонент системы базы данных, находящийся справа, - сама база данных.

В общем случае, можно сказать, что ***база данных*** - это самодокументированное собрание интегрированных записей. Для всех реляционных баз данных (а это сегодня почти все базы данных) это определение можно модифицировать так: ***база данных*** - это самодокументированное собрание связанных таблиц.

Под ***самодокументированностью*** подразумевается, что описание структуры базы данных содержится в самой базе данных. Благодаря этому всегда можно узнать содержимое базы данных, просто посмотрев на нее. Данные о структуре базы данных называются ***метаданными***.

Примерами метаданных служат имена таблиц, имена столбцов и таблицы, которым они принадлежат, а также свойства таблиц и столбцов, и так далее

Описания баз данных относятся к метаинформации, т. е. информации об информации. Описание баз данных часто называют ***схемой***. Кроме того, в БнД могут присутствовать описания отдельных частей баз данных с точки зрения отдельных пользователей — ***подсхемы.***

Кроме описания баз данных в состав метаинформации, хранимой в БнД, может включаться информация о предметной области, необходимая для проектирования системы, о пользователя БнД, о проектных решениях и некоторая другая информация.

Централизованное хранилище метаинформации называется ***словарём*** ***данных***. В литературе используются также термины словарь-справочник, энциклопедия, репозиторий. В некоторых источниках выявляются различия между ними, в других они используются как синонимы. Для данного уровня рассмотрения для нас эти различия несущественны.

Роль словарной системы особенно возрастает при использовании средств автоматизированного проектирования информационных систем. Для большинства из них репозиторий является ядром всей системы

Кроме того, база данных имеет ***индексы*** и другие структуры, которые существуют для облегчения представления баз данных. Индекс похож на предметный указатель в конце книги и показывает, где искать определенные записи в таблице.

***Хранимая процедура*** - это программа, которая хранится в базе данных. Некоторые хранимые процедуры являются утилитами для базы данных.

***Триггер*** - это процедура, которая выполняется при воникновении определенных ситуаций в данных.

Наконец, некоторые базы данных содержат ***метаданные приложений*** - простые данные, которые описывают элементы приложения, такие как формы и отчеты. Например, Microsoft Access содержит метаданные приложения как часть своих баз данных.

**7. Структурирование данных**

Создавая базу данных, пользователь стремится упорядочить информацию по различным признакам и быстро извлекать выборку с произвольным сочетанием признаков. Сде­лать это возможно, только если данные структурированы.

***Структурирование*** — это введение соглашений о способах представления данных.

Неструктурированными называют данные, записанные, например, в текстовом файле.

Личное дело № 16493, Сергеев Петр Михайлович, дата рождения 1 января 1976 г.; Л/д № 16593. Петрова Анна Владимировна, дата рожд. 15 марта 1975 г.; № личн. дела 16693. д.р. 14.04.76. Анохин Андрей Борисович.

Рисунок 3 - Пример неструктурированных данных

Чтобы автоматизировать поиск и систематизировать эти данные, необходимо выработать определенные соглашения о способах представления данных, т.е. дату рождения нужно записывать одинаково для каждого студента, она должна иметь одинаковую длину и определенное место среди остальной информации. Эти же замечания справедливы и для остальных данных (номер личного дела, фамилия, имя, отчество).

**Пример**. После проведения несложной структуризации с информацией, указанной на рисунке 3, она будет выглядеть так, как это показано на рисунке 4.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **№ личного дела** | **Фамилия** | **Имя** | **Отчество** | **Дата рождения** |
| 11123 | Агеева | Ирина | Сергеевна | 01.05.1974 |
| 11168 | Байжанов | Омар | Серикович | 02.03.1982 |
| 11125 | Сармурзин | Аслан | Кожахметович | 05.06.1984 |

Рисунок 4 - Пример структурированных данных

Всвязи с вышесказанным, можно дать еще одно определение базы данных.

***База данных*** (БД) — это поименованная совокупность структурированных данных, относящихся к определенной предметной области.

**8. Структурные элементы базы данных.**

Понятие базы данных тесно связано с такими понятиями структурных элементов, как поле, запись, файл (таблица) (рисунок 5).

***Поле*** *—* элементарная единица логической организации данных, которая соответствует неделимой единице информации — реквизиту.

Для описания поля используются следующие *характеристики:*

***имя****,* например. Фамилия, Имя, Отчество, Дата рождения;

***тип****,* например, символьный, числовой, календарный;

***длина****,* например, 15 байт, причем будет определяться максимально возможным количеством символов;

***точность*** для числовых данных, например два десятичных знака для отображения дробной части числа.

*Поле*

|  |  |  |  |
| --- | --- | --- | --- |
| Имя поля 1 | Имя поля 2 | Имя поля 3 | Имя поля 4 |
| Запись |  |  |  |
|  |  |  |  |
|  |  |  |  |

##### Рисунок 5 - Основные структурные элементы БД

***Запись*** — совокупность логически связанных полей. Экземпляр записи — отдельная реализация записи, содержащая конкретные значения ее полей.

***Файл (таблица)*** — совокупность экземпляров записей одной структуры.

Описание логической структуры записи файла содержит последовательность расположения полей записи и их основные характеристики, как это показано на рисунке 6.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Имя файла | | | | | |
| Поле | | Признак ключа | Формат поля | | |
| Имя (обозначение) | Полное наименование | Тип | Длина | Точность (для чисел) |
| Имя 1 |  |  |  |  |  |
| …. |  |  |  |  |  |
| …. |  |  |  |  |  |
| Имя n |  |  |  |  |  |

Рисунок 6 - Описание логической структуры записи файла

В структуре записи файла указываются поля, значения которых являются ***ключами****:*

***первичными*** (ПК), которые идентифицируют экземпляр записи, и ***вторичными***(ВК), которые выполняют роль поисковых или группировочных признаков (по значению вторичного ключа можно найти несколько записей).

**Пример.** На рисунке 7 приведен пример описания логической структуры записи файла (таблицы) СТУДЕНТ. Структура записи файла СТУДЕНТ линейная, она содержит записи фиксированной длины. Повторяющиеся группы значений полей в записи отсутствуют. Обращение к значению поля производится по его номеру.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Имя файла: Студент | | | | | |
| Поле | | Признак ключа | Формат поля | | |
| Имя (обозначение) | Полное наименование | Тип | Длина | Точность (для чисел) |
| Номер | № личного дела | \* | Симв | 5 |  |
| Фамилия | Фамилия студента |  | Симв | 15 |  |
| Имя | Имя студента |  | Симв | 10 |  |
| Отчество | Отчество студента |  | Симв | 15 |  |
| Дата | Дата рождения |  | Дата | 8 |  |

Рисунок 7 - Описание логической структуры записи файла СТУДЕНТ

**9. Уровни представления данных**

Архитектура системы баз данных предусматривает следующие уровни: уровень внешних моделей, концептуальный уровень и физический уровень (рисунок 8).

Трехуровневая модель системы управления базами данных предложена американским комитетом по стандартизации ANSI. Эта архитектура позволяет обеспечить логическую (между уровнями 1 и 2) и физическую (между уровнями 2 и 3) независимость при работе с данными. Логическая независимость предполагает возможность изменения одного приложения без корректировки других приложений, работающих с этой же базой данных. Физическая независимость предполагает возможность переноса хранимой информации с одних носителей на другие при сохранении работоспособности всех приложений БД

Внешняя модель данных 1

Внешняя модель данных 2

Внешняя модель данных 2

Внешняя модель данных n

Концептуальный уровень

База данных

*Уровень 1*

*Уровень 2*

*Уровень 3*

##### Рисунок 8 - Уровни представления данных

1. Уровень внешних моделей. Этот уровень определяет точку зрения на БД отдельных приложений. Каждое приложение «видит» и обрабатывает только те данные, которые необходимы именно этому приложению.
2. Концептуальный уровень – центральное управляющее звено. Здесь база данных представлена в наиболее общем виде, который» объединяет данные, используемые всеми приложениями, работающими с данной БД. Практически концептуальный уровень отражает обобщенную модель предметной области (объектов реального мира), для которой создавалась база данных.
3. Физический уровень – собственно данные, расположенные в файлах или в страничных структурах на внешних носителях информации.

**Контрольные вопросы:**

1. Что такое база данных, система управления базами данных, банк данных?
2. Перечислите основные операции с базами данных.
3. Чем отличается банк данных от системы управления базами данных?
4. Перечислите преимущества банков данных.
5. Какие требования предъявляются к банкам данных?
6. Перечислите и охарактеризуйте компоненты системы баз данных.
7. В чем заключаются функции прикладных программ?
8. Охарактеризуйте функции СУБД.
9. Что представляют из себя языковые средства СУБД? Какие принципы используют при их проектировании?
10. Перечислите компоненты базы данных и дайте их определения.

Дайте определение структурным элементам базы данных.