

Система управления базой данных

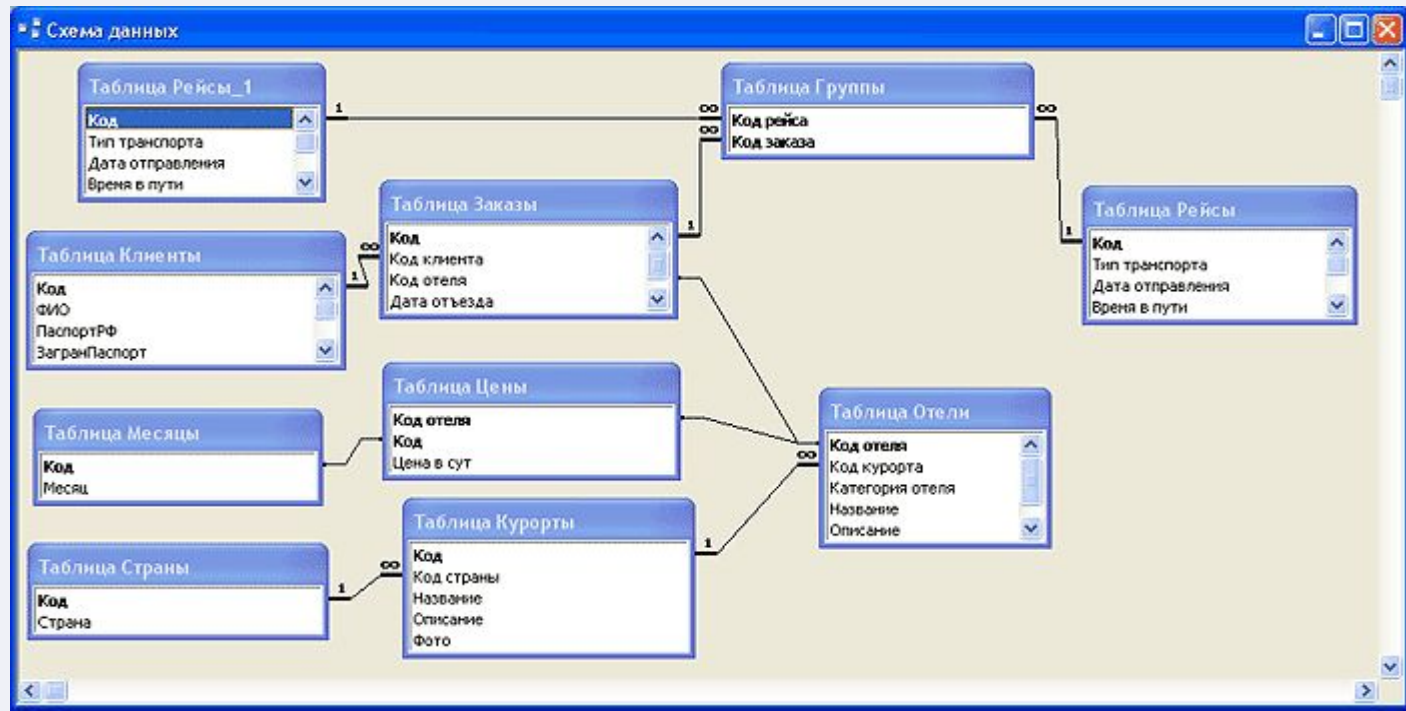


Цель любой информационной системы – обработка данных об объектах реального мира. Часть реального мира, подлежащая изучению для организации управления и автоматизации, называется предметной областью

Базы данных всегда были важнейшей темой при изучении информационных систем. Однако в последние годы всплеск популярности Интернета и бурное развитие новых технологий для Интернета сделали знание технологии баз данных для многих одним из актуальнейших путей карьеры. В то же время Интернет-технология обеспечивает пользователям стандартизированные и доступные средства публикации содержимого баз данных.

С самого начала развития вычислительной техники образовались два основных направления ее использования.

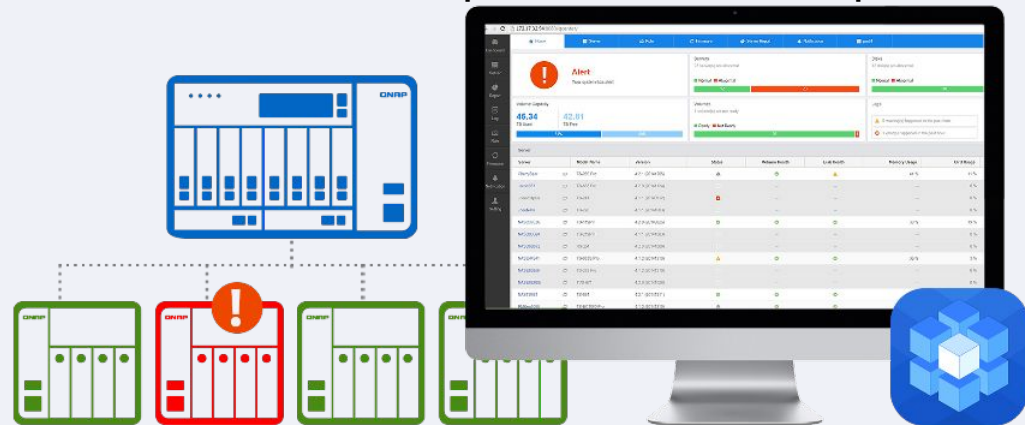
Первое направление - применение вычислительной техники для выполнения численных расчетов, которые слишком долго или вообще невозможно производить вручную. Становление этого направления способствовало интенсификации методов численного решения сложных математических задач, развитию класса языков программирования, ориентированных на удобную запись численных алгоритмов, становлению обратной связи с разработчиками новых архитектур ЭВМ.



Второе направление, это использование средств вычислительной техники в автоматических или автоматизированных информационных системах. В самом широком смысле информационная система представляет собой программный комплекс, функции которого состоят в поддержке надежного хранения информации в памяти компьютера, выполнении специфических для данного приложения преобразований информации и/или вычислений, предоставлении пользователям удобного и легко осваиваемого интерфейса. Обычно объемы информации, с которыми приходится иметь дело таким системам, достаточно велики, а сама информация имеет достаточно сложную структуру.

Историческим шагом явился переход к использованию централизованных систем управления файлами. С точки зрения прикладной программы *файл* - это именованная область внешней памяти, в которую можно записывать и из которой можно считывать данные. Правила именования файлов, способ доступа к данным, хранящимся в файле, и структура этих данных зависят от конкретной системы управления файлами и, возможно, от типа файла. Система управления файлами берет на себя распределение внешней памяти, отображение имен файлов в соответствующие адреса во внешней памяти и обеспечение доступа к данным.

Все современные файловые системы поддерживают многоуровневое именование файлов за счет поддержания во внешней памяти дополнительных файлов со специальной структурой - каталогов. Каждый каталог содержит имена каталогов и/или файлов, содержащихся в данном каталоге. Таким образом, полное имя файла состоит из списка имен каталогов плюс имя файла в каталоге, непосредственно содержащем данный файл.



Взять данные из гороскопа:

18.06.1942 12:00 Натальный - Пол Маккартни

Фамилия, Имя, Отчество:

Пол Маккартни

Событие:

Космограмма Рождения

Место события:

Ливерпуль

Число: 18 Четверг Час: 12
Месяц: 06 Июнь Мин: 00
Год: 1942 Не високосный Сек: 00

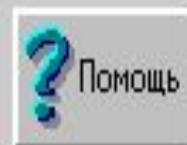
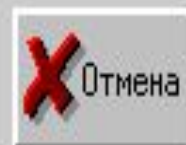
Разница с Гринвичем: 0 : 00 : 00 ?

Широта: 44 ° 05 ' 00 " N S
Долгота: 000 ° 00 ' 00 " E W ?

Запомнить в каталог:

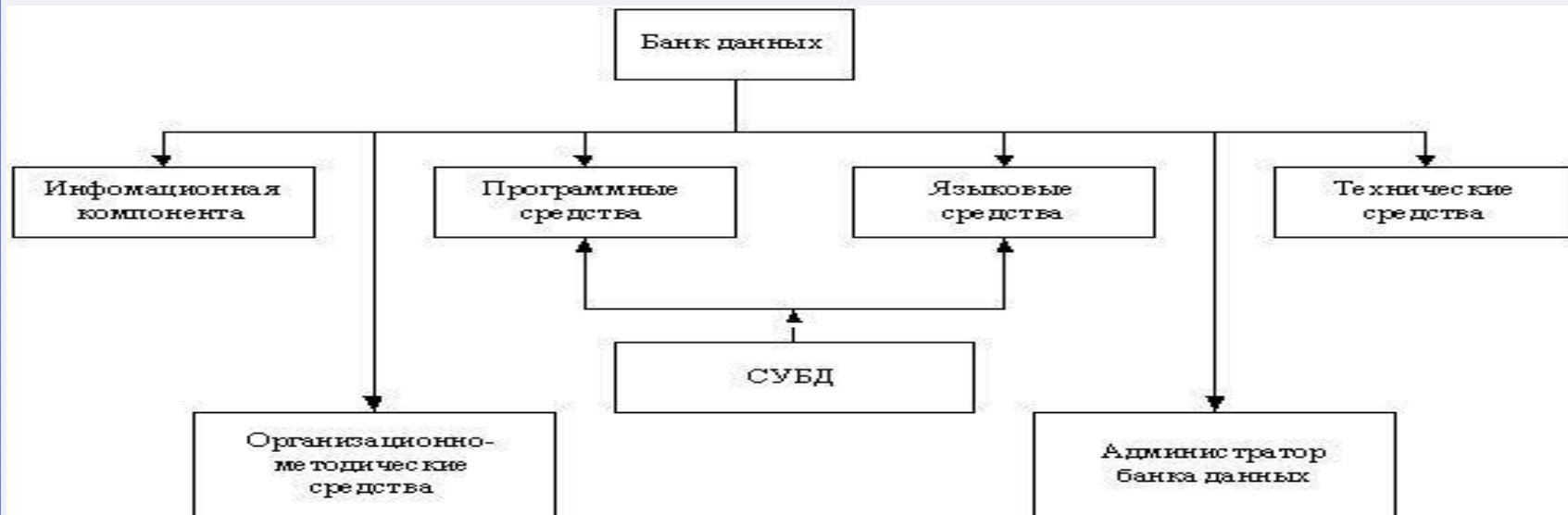
Путь к каталогу в банке:

Знаменитые Люди



Компонентами банка данных являются:

- база данных;
- система управления базой данных;
- вычислительная система (операционная система и технические средства);
- администратор базы данных (это группа специалистов, без которых невозможно функционирование и развитие базы данных);
- словарь данных;
- обслуживающий персонал.



По условиям предоставления услуг различают **бесплатные и платные банки данных**. Платные БД, в свою очередь, делятся на неприбыльные и коммерческие. Неприбыльные банки данных функционируют на принципе самоокупаемости и не ставят своей целью получение прибыли. Это обычно БД социально значимой информации, имеющей широкий круг пользователей, или научной, библиотечной информации. Основной целью создания коммерческих банков данных является получение прибыли от информационной деятельности.

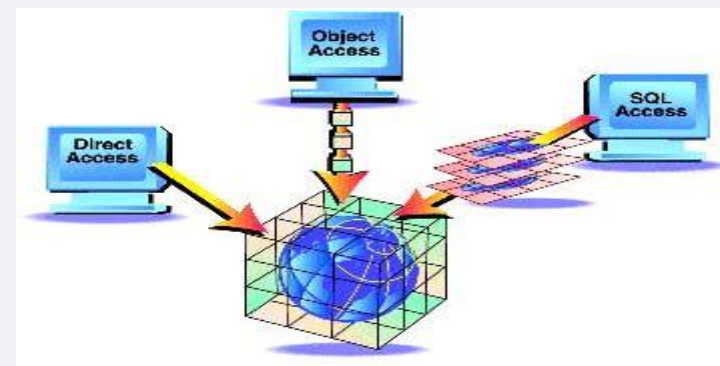
По форме собственности банки данных делятся на государственные и негосударственные.

По степени доступности различают общедоступные и с ограниченным кругом пользователей.

База данных – это информационная модель предметной области. Это – ядро банка данных. База данных – это совокупность взаимосвязанных, хранящихся вместе данных при наличии такой минимальной избыточности, которая допускает их использование оптимальным образом для одного или нескольких приложений. Это совокупность специальным образом организованных данных (файлов), хранимых во внешней памяти и используемых в качестве входной информации для решения задач.

СУБД – это программа, с помощью которой реализуется централизованное управление данными, хранимыми в базе, доступ к ним, поддержка их в необходимом состоянии. Система управления базами данных управляет данными во внешней памяти, обеспечивает надежное хранение данных и поддержку соответствующих языков базы данных.

СУБД используется для создания самой базы данных, таблиц и других поддерживаемых структур, чтение и изменение данных в базе данных. Для этого СУБД получает запросы SQL (или какие-нибудь другие запросы) и преобразует их в действия по отношению к базе данных. Другая функция СУБД — поддержка всех структур базы данных. СУБД контролирует работу, следя, чтобы изменения одного пользователя не пересекались с изменениями другого. Кроме того, СУБД содержит систему безопасности, которая используется для проверки того, что только авторизованные пользователи выполняют определенные действия с базой данных. Пользователям могут не позволить знакомиться некоторыми данными, а также ограничить их действия только определенными изменениями в определенных данных. СУБД обеспечивает возможность резервного копирования данных из базы данных и восстановления их в случае необходимости.



Одним из основных требований к СУБД является надежность хранения данных во внешней памяти. Под надежностью хранения понимается то, что СУБД должна быть в состоянии восстановить последнее согласованное состояние БД после любого аппаратного или программного сбоя.

Любое современное учреждение информационно поддерживается профессиональной СУБД. Эволюция систем управления базами данных прошла путь от систем, опиравшихся на иерархическую и сетевую модели данных, до систем третьего и четвертого поколений, для которых характерны идеи объектно-ориентированного подхода к организации данных и принципы распределенной обработки данных.

Системы управления базами данных можно классифицировать по способу установления связей между данными, по характеру выполняемых ими функций, сфере применения, по числу поддерживаемых моделей данных, по характеру используемого языка общения с базой данных и другим параметрам.

По выполняемым функциям СУБД подразделяются на операционные и информационные. По сфере применения СУБД подразделяются на универсальные и проблемно-ориентированные. По используемому языку общения СУБД подразделяются на замкнутые, имеющие собственные самостоятельные языки общения пользователей с базами данных, и открытые, в которых для общения с базой данных используется язык программирования, расширенный операторами языка манипулирования данными.

По числу поддерживаемых уровней моделей данных СУБД подразделяются на одно-, двух-, трехуровневые системы.

По способу установления связей между данными различают реляционные, иерархические и сетевые базы данных.

По способу организации хранения данных и выполнения функций обработки базы данных подразделяются на централизованные и распределенные.

В *централизованной базе данных* обеспечивается простота управления, улучшенное использование данных на местах при выполнении дистанционных запросов, более высокая степень одновременности обработки, меньшие затраты на обработку.

Распределенная база данных предполагает хранение и выполнение функций управления данными в нескольких узлах и передачу данных между этими узлами в процессе выполнения запросов. В такой базе данных не только различные ее таблицы могут храниться на разных компьютерах, но и разные фрагменты одной таблицы так же. При этом для пользователя не имеет значения, как организовано хранение данных, он работает с такой базой как с централизованной.

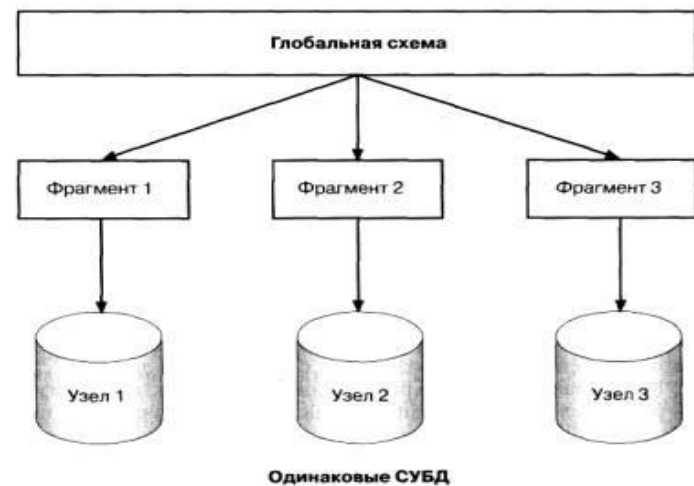


Рис. 2.1. Архитектура однородной распределенной базы данных

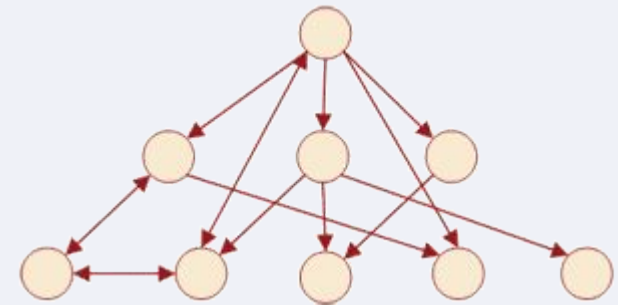
Модель данных – это совокупность структур данных и операций их обработки.

СУБД основывается на использовании иерархической, сетевой и реляционной модели и их комбинации.

Иерархическая модель данных – объекты связаны иерархическими отношениями, которые образуют ориентированный граф (перевернутое дерево). Основные понятия: уровень, узел, связь. Узел – вершина графа.

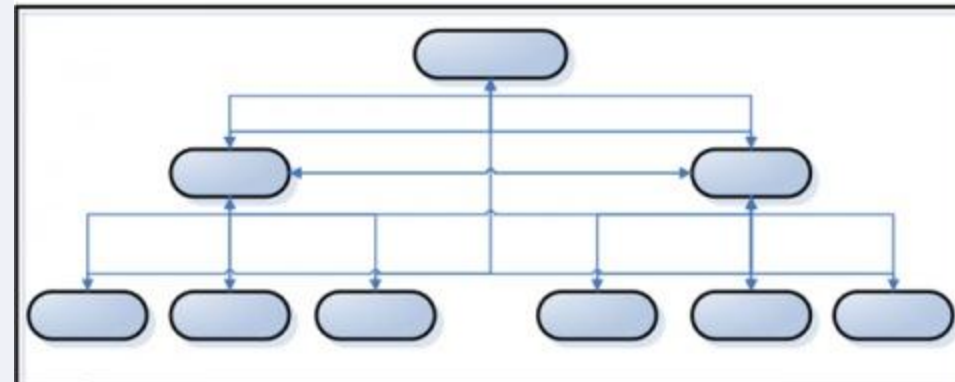
К каждой записи базы данных существует только один (иерархический) путь от корневой записи.

Сетевая модель данных – каждый элемент может быть связан с любым другим.



Реляционная модель данных – ориентирована на организацию данных в виде двумерных таблиц.

Структура реляционной модели БД



Уровни представления данных

База данных и программные средства для их создания и ведения (СУБД) имеют многоуровневую архитектуру. Различают концептуальный, внутренний и внешний уровни представления данных, которым соответствуют модели аналогичного назначения.

Концептуальный уровень - соответствует логическому аспекту представления данных.

Внутренний уровень отображает требуемую организацию данных в среде хранения и соответствует физическому аспекту представления данных (внешние носители).

Внешний уровень поддерживает частные представления данных, требуемые конкретным пользователем (задаются экранные формы, доступ к данным, режимы обработки).

Проектирование базы данных состоит в построении комплекса взаимосвязанных моделей данных. Важнейшим этапом проектирования базы данных является разработка инфологической (информационно-логической) модели предметной области. Информационно-логическая модель предметной области отражает предметную область в виде совокупностей информационных объектов и их структурных связей.

Для работы с базами данных используются **специальные языки**, в целом называемые языками баз данных.

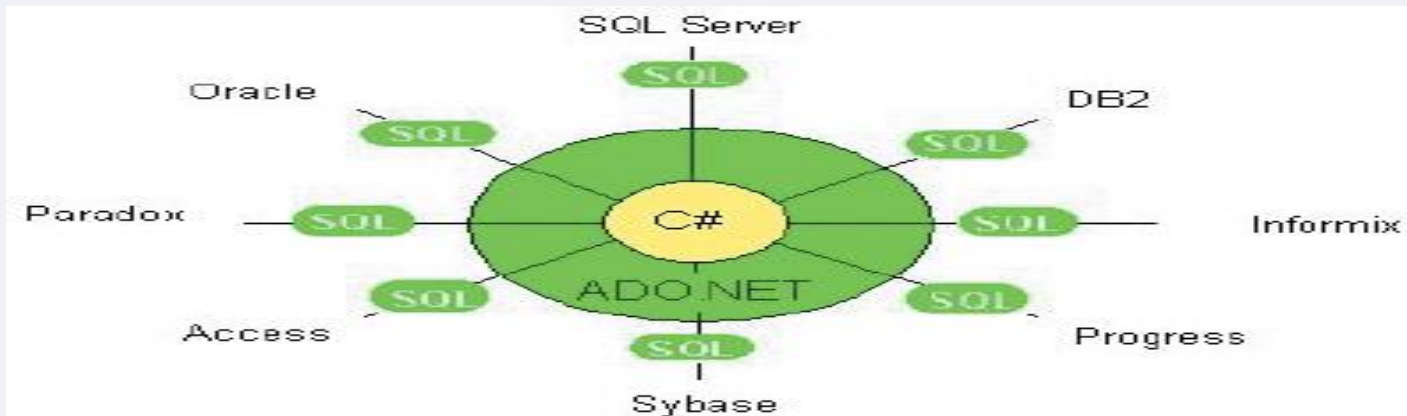
В современных СУБД обычно поддерживается единый интегрированный язык, содержащий все необходимые средства для работы с БД, начиная от ее создания, и обеспечивающий базовый пользовательский интерфейс с базами данных.

Стандартным языком наиболее распространенных в настоящее время реляционных СУБД является язык **SQL** (Structured Query Language). Язык SQL является международным стандартом языка запросов. Он представлен в явной синтаксической форме и является инструментом, предназначенным для обработки и чтения данных, содержащихся в компьютерной базе данных. Это язык программирования, который применяется для взаимодействия пользователя с базой данных. Если пользователю необходимо прочитать данные из базы данных, он запрашивает их у СУБД с помощью языка SQL.



Язык SQL позволяет определять схему реляционной БД и манипулировать данными. Специальные операторы языка SQL позволяют определять так называемые представления БД, фактически являющиеся хранимыми в БД запросами (результатом любого запроса к реляционной БД является таблица) с именованными столбцами. Авторизация доступа к объектам БД производится также на основе специального набора операторов SQL.

В отличие от языка SQL, заданного в явной форме, широко используются **интерфейсы с неявным заданием языка**. В этом случае функции языков могут быть доступны косвенным образом, когда они реализуются в форме различного рода меню, диалоговых сценариев или заполняемых пользователем таблиц. По таким входным данным интерфейсные средства формируют адекватные синтаксические конструкции языка интерфейса и передают их на исполнение. Интерфейсы с неявным использованием языка широко используются в СУБД для ПЭВМ. Примером такого языка является язык QBE. (Query By Example).



Объектно-ориентированное программирование начало развиваться в середине восьмидесятых годов прошлого века и привело к созданию объектно-ориентированных СУБД. Однако наиболее активно это направление развивается в последние годы. Возникновение направления ООБД определяется прежде всего потребностями практики: необходимостью разработки сложных информационных прикладных систем, для которых технология предшествующих систем БД не была вполне удовлетворительной. Целью этих продуктов была способность хранить объекты из объектно-ориентированного программирования (например, из языков С++ или Java) в базе данных, не преобразуя их в реляционный формат. Были разработаны объектно-ориентированные СУБД, такие как Oracle 8i и 9i, что позволило создавать как реляционное, так и объектное представление данных одной и той же базы данных.

Компьютерные базы данных

Большая часть публикуемой информации представлена также в удобном формате в виде компьютерных баз данных, предназначенных для распространения в электронном виде. В 1980—90 годы количество баз данных, равно как и их поставщиков, резко возросло.

Процесс поиска данных отличается доступностью, быстротой и простотой. Предоставляется возможность доступа к сотням наименований баз данных. При этом обеспечивается практически мгновенный доступ к требуемой информации благодаря упрощенному процессу поиска, для которого используются стандартные, одинаковые у всех поставщиков поисковые протоколы и команды.

Компьютерные базы данных делятся на базы данных с доступом в режиме online, offline и через Internet.

Базы данных с доступом в *режиме online* (online databases) хранятся в центральном банке данных. Доступ к ним осуществляется посредством компьютера (или иного терминала) через телекоммуникационную сеть.

Доступ, поиск и анализ Internet-баз данных (internet databases) осуществляется посредством Internet. Сведения из них можно загружать и сохранять на компьютере или вспомогательном запоминающем устройстве.

Базы данных с доступом в *режиме offline* (offline databases) представляют собой информацию, хранящуюся на дискетах или компакт-дисках и доступную для потребителей без использования внешней телекоммуникационной сети.

Базы данных в свою очередь подразделяются на *библиографические, цифровые, текстовые, справочные и специализированные.*

Библиографические базы данных (bibliographic databases) состоят из ссылок и цитат из статей журналов, газет, отчетной документации маркетинговых исследований, технических докладов, правительственной документации и т.п. Они часто содержат краткие изложения и отрывки из цитируемых материалов.

Цифровые базы данных (numeric databases) содержат цифровую и статистическую информацию. Например, некоторые цифровые базы данных предоставляют хронологически систематизированную информацию о структуре и специфической продукции ряда компаний.

Текстовые базы данных (full-text databases) состоят из полных текстов оригинальных документов.

Таким образом, системы управления базами данных (СУБД) - это очень важная тема, без основательного знакомства с которой в наше время невозможно быть не только квалифицированным программистом, но даже и грамотным пользователем компьютеров.