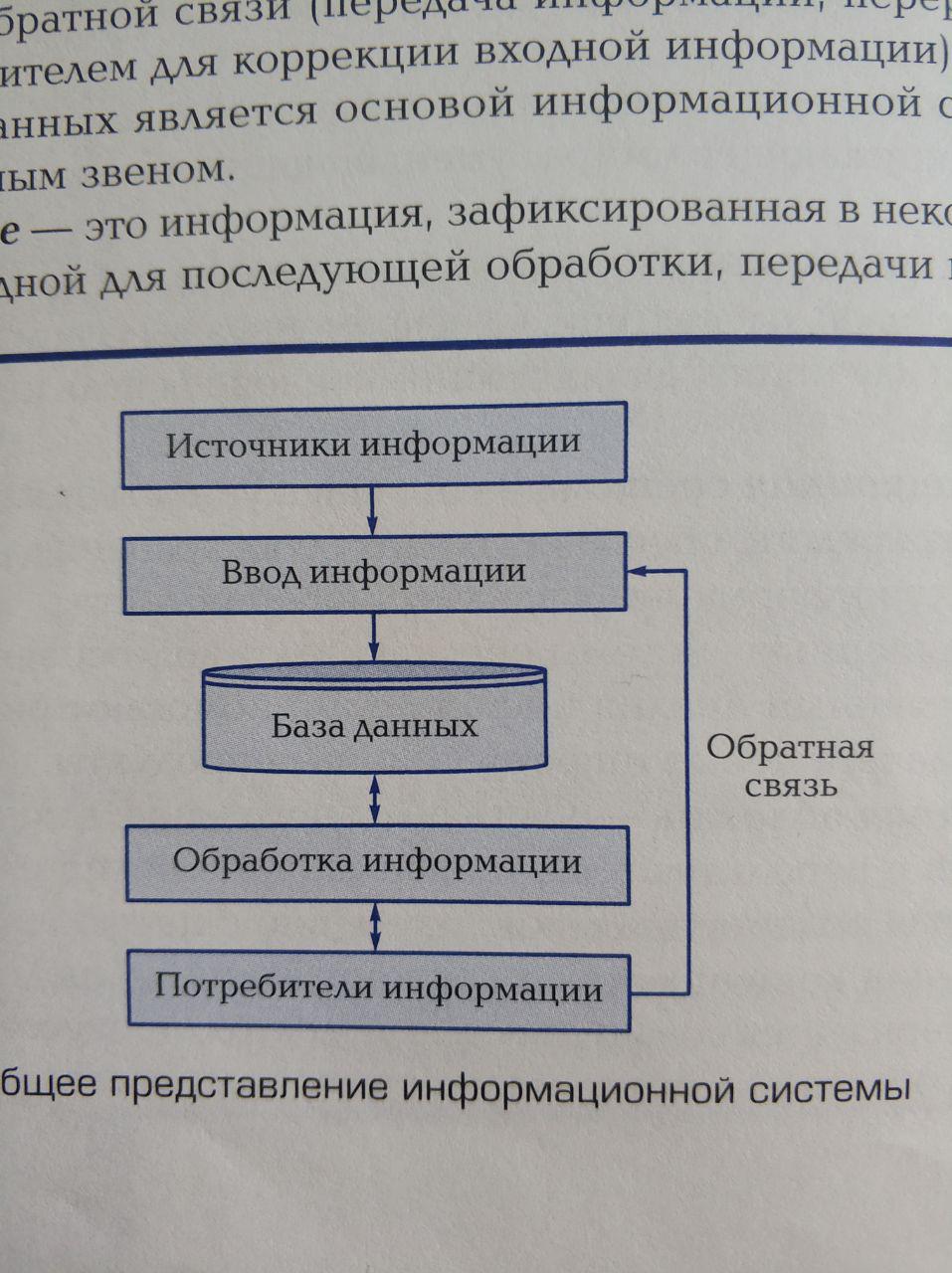
1. Базы данных и информационные системы. Основные определения.

Информация - это сведения об объектах, явлениях, процессах, событиях окружающего мира уменьшающие степень неопределенности знаний о них. Своевременная выдача достоверной информации для принятия решений это основная цель информационных систем. Информационная система - это совокупность технических и программных средств, обеспечивающих сбор, хранение, обработку, поиск, выдачу информации в задачах любой области. Предметная область - часть реального мира, данные о которой хранятся и используются в информационной системе. Предметная область подлежит изучению с целью организации управления и автоматизации. Она характеризуется совокупностью объектов, использующих их процессов и множеством пользователей. Ее анализ предшествует созданию любой информационной системы.

Информационный объект - это описание некоторой сущности предметной области.

Информационный объект можно представить в виде схемы.



Данные - это информация, зафиксированная в некоторой форме, пригодная для дальнейшей машиной обработки, передачи и хранения. База данных - это именованная совокупность взаимосвязанных данных отображающая состояние объектов в некоторой предметной области. Структурой данных называют совокупность правил и ограничений, которые отображают связи между отдельными частями данных. Обработка данных - совокупность задач для преобразования массивов данных. Система обработки данных - набор аппаратных и программных средств, осуществляющих выполнение задач по управлению данными. Управление данными - это круг операций с ними, которые необходимы для успешного функционирования системы обработки данных. Метаданные - это данные о данных, это описание собственной структуры базы данных.

2. Системы управления базами данных.

Это совокупность языковых и программных средств, предназначенных для управления, создания и использования баз данных. Они являются инструментальными средствами для извлечения данных, их обработки, изменения и анализа. СУБД предоставляет удобный, быстрый и контролируемый доступ к данным, обладает средствами для обеспечения целостности данных, поддержания баз данных в рабочих состояниях, обеспечения секретности, восстановления и сохранности информации в многопользовательском окружении.

СУБД предназначена в основном для профессиональных разработчиков. Обычный пользователь работает с базами данных с помощью специальных приложений, которые могут создаваться как внутри СУБД, так и вне ее с помощью систем программирования. СУБД должна обеспечить параллельную независимую работу таких приложений с единой базой данных.

Более современной формой организации хранения и доступа к информации является банк данных. Банк данных - система специально организованных данных, включая базы данных, программных, языковых, организационно-методических средств для обеспечения централизованного накопления и коллективного многоцелевого использования.

Словарь данных - это подсистема банка данных предназначенная для централизованного хранения информации о структурах данных, взаимосвязях файлы базы данных друг с другом и т.п.

3. Распределенные базы данных. Базы знаний.

Существуют базы данных, в которых предполагается хранение данных распределенное в компьютерной сети. Распределенная база данных – это база данных, включающая фрагменты из нескольких баз данных, которые располагаются на различных узлах компьютерной сети и возможно управляются различными СУБД. С точки зрения пользователей и прикладных программистов она выглядит как обычная локальная база данных. Распределенная структура повышает возможности совместного использования удаленных данных, надежность, доступность и производительность системы, улучшить ее масштабируемость.

Современные информационные системы зачастую используют базы знаний. Знания – выявленные закономерности предметной области, позволяющие решать задачи этой области. База знаний – совокупность знаний предметной области. Системы, использующие базы знаний называются интеллектуальными.

4. Основные категории пользователей и разработчиков баз данных.

Основная категория пользователей, для которых создается база данных – это конечные пользователи. Ими могут быть случайные одноразовые пользователи, регулярные пользователи. От конечных пользователей не должно требоваться специальных знаний в области вычислительной техники и программирования.

Разработчики и администраторы приложений. Эта группа пользователей функционирует во время проектирования, создания и реорганизации базы данных. Администраторы приложений координируют работу разработчиков. Разработчики конкретных приложений работают с той частью данных из базы, которая требуется для этих приложений. Наиболее сложные обязанности возложены на группу администратора баз данных – это группа пользователей, которые на начальной стадии разработки базы данных отвечает за ее оптимальную организацию и одновременную работу конечных пользователей. На стадии развития и реорганизации они отвечают за возможность корректной реорганизации базы данных без нарушения текущей ее эксплуатации.

В составе группы администратора базы данных должны быть:

1. Системные аналитики.

2. Проектировщики структур данных и внешнего информационного обеспечения.

3. Проектировщики технологических процессов обработки данных.

4. Системные и прикладные программисты.

5. Операторы и специалисты по техническому обслуживанию.

5. Основные функции СУБД.

Изначально системы управления базами данных пытались построить по принципам файловых систем, но получившаяся СУБД оказалась громоздкой и неэффективной. Поэтому были разработаны системы обработки данных. Они включили в себя сами данные, СУБД и прикладное ПО, которое обращается к данным через СУБД. Общими функциями СУБД являются:

1. Управление данными во внешней памяти.

2 Управление данными в ОП.

3. Управление транзакциями.

4. Журнализация, резервное копирование и восстановление.

5. Поддержка языков баз данных.

6. Управление данными во внешней памяти.

СУБД должна представлять пользователям возможность:

Сохранять извлекать и обновлять данные в базу данных - это самая фундаментальная функция назначения СУБД. Способ реализации этой функции должен быть скрыт от конечного пользователя.

Контролировать доступ к данным – это возможность обеспечить только санкционированной доступ к базе данных, используя поддержку уровня доступа к БД и отдельным её элементам. Каждый пользователь должен иметь доступ только к тем данным, которые доступны для него в соответствии с его пользовательскими правами.

Обеспечивать параллельную работу нескольких пользователей – СУБД имеют механизмы, которые гарантируют корректное выполнение данных многими пользователями при одновременном доступе.

Поддерживать целостность данных – осуществляется инструментальными средствами контроля для того чтоб данные и их изменения соответствии заданным правилам. Целостность базы данных это её свойство, означающее, что в ней задерживается полное не противоречие и адекватное отражающее требуемую область информации. Целостное состояние БД описывается с помощью ограничений целостности в виде условий, которым должны удовлетворять данные.

7. Управление данными в ОЗУ.

Обычно размер базы данных значительно больше объема оперативной памяти. Если при обращении к любым данным будет производиться обмен с внешней памятью, то скорость работы СУБД будет равно скорости работы внешней памяти. Единственным реальным способом увеличения этой скорости является буферизация оперативной памяти.

Буфер это область оперативной памяти, в которых временно хранятся фрагменты базы данных, данные из которых предлагается использовать, при обращении к СУБД или записать в базу после обработки.

8. Управление транзакциями.

Транзакции это число действий над базами данных, которые последовательно выполняются по принципу все или не одного. Если транзакции выполняются успешно, то СУБД фиксирует произведенные действия БД во внешней памяти. Если хотя бы одно действие транзакций не выполняется, то все изменения произведенные ей в БД отменяются. Это необходимо для поддержания логической целостности БД.

При параллельном выполнении транзакций возможно возникновение конфликтов, разрешение которых является функцией СУБД, при обнаружении таких случаев обычно производится отмена изменений, произведенных одной или нескольких транзакций - **откат транзакций**.

9. Журнализация, резервное копирование и восстановление.

Одним из основных требований к СУБД является надёжность хранения данных во внешней памяти, в том числе защита их физической и логической целостности. Это значит, что СУБД должна представлять возможность восстановить последнее согласованное состояние БД после любого сбоя. Защита физической целостности включает в себя журнализацию изменений, резервное копирование и восстановление БД. **Журнализация изменений** - это последовательная запись во внешнюю память всех изменений, выполняемых в БД: порядковый номер, тип и время изменения; идентификатор транзакций, объект, подвергшийся изменению, предыдущее и новое состояние блокнота.

Журнал содержит отметки начала и завершения транзакций и отметки принятия контрольной точки, он не доступен пользователями СУБД. Для повышения надёжности могут поддерживаться 2 копии журнала.

**Резервное копирования базы данных** – это процесс копирования базы данных на носителе, предназначенное для восстановления данных, в оригинальном или в новом месте их расположения, в случае повреждения или разрушения.

**Восстановление базы данных** – это функция СУБД, которая после логических или физических сбоев приводит базу данных в актуальное состояние.

10. Поддержка языков баз данных.

В современных СУБД обычно поддерживается единый интегрированный язык, содержащий средства для работы с БД и обеспечивающий базовый пользовательский интерфейс с базами данных. Стандартным языком реляционной СУБД в настоящее время является SQL. Кроме своих основных функций СУБД предоставляет вспомогательные утилиты для эффективного администрирования базы данных:

1. Отвечает за экспорт и импорт данных.

2. **Мониторинг базы данных** – отслеживание характеристик, функционирования и использование БД.

3. **Статистический анализ** – для оценки производительности или степень использования БД.

4. Реорганизация индекса

5. Сборка не используемых записей и перераспределение памяти для физического устранения удаленных записей с внешних носителей, дефрагментация освобождённого пространства.

В современных СУБД можно выделить ядро СУБД, процессор языка, БД, подсистему поддержки выполнения, внешние утилиты для выполнения возможности БД.

Ядро СУБД это набор программных модулей необходимый и достаточный для создания и поддержания БД. Сервисные программы представляют дополнительные услуги, зависящие от конкретной предметной области и потребности конкретного пользователя. Процессор языка БД компилирует операторы языка в программный машинный код.

11. Архитектура баз данных.

В 1975 году на заседании Американского Национального Института стандартов была предложена обобщенная трехуровневая модель архитектуры СУБД. Она включала внешний, концептуальный и физический уровни. Основная цель этой идеи заключается в отделении пользовательского представления базы данных от ее физического представления.

**Уровень внешних моделей** – это самый верхний уровень, где каждая модель имеет свое представление данных. Отдельные группы пользователей работают только с теми данными, к которым они имеют доступ. При рассмотрении приложений, работающих с одной базой данных, предполагается, что они могут работать параллельно и независимо друг от друга. Именно СУДБ должна обеспечить такую возможность.

**Концептуальный уровень** архитектуры является основным и служит для представления базы данных в общем виде для всех ее приложений и не зависит от них. Концептуальный уровень является промежуточным уровнем в трехуровневой архитектуре и обеспечивает представление всей информации БД в абстрактной форме – это полное представление требований к данным предметной области, независящее от способов их представления и хранения. Описание БД на этом уровне называется концептуальной схемой. Концептуальная схема включает объекты и их атрибуты, связи между объектами, ограничения на данные и т.д.

**Физический уровень** поддерживает представление базы данных в среде хранения. Описание БД на этом уровне называется внутренней схемой или схемой хранения. **Физический уровень** – это собственно данные, расположенные в файлах или страничных структурах, расположенных на внешних носителях. На этом уровне необходимо достичь оптимальной производительности и обеспечить экономное использование дискового пространства. На физическом уровне хранится информация о распределении дискового пространства для хранения данных и индексов, описание подробностей хранения записей, сведения о размещении записей, сведения о сжатии данных и выбранных методах шифрования. Физический уровень контролируется ОС-ой под руководством СУБД. В соответствии с междууровневыми отображениями принято выделять логическую и физическую независимость данных. Логическая независимость означает защищенность внешних схем от изменений, вносимых в концептуальную схему, кроме того, она предполагает возможность изменения одного приложения без изменения других приложений. Физическая независимость означает защищенность концептуальной схемы от изменений, вносимых в схему хранения. Также она предполагает возможность переноса хранимой информации с одних носителей на другие.

12. Понятие модели данных.

В классической теории БД **модель данных** – это формальная теория представления и обработки данных в СУБД. Она включает в себя 3 аспекта:

1. Аспект структуры – методы описания типов и логических структур данных в БД.

2. Аспект манипуляции – методы манипулирования данными.

3. Аспект целостности – методы описания и поддержки целостности БД.

Модель данных описывает информационные объекты предметной области, взаимосвязи между ними и позволяет: