



ИНФОРМАЦИОННЫЙ ПРАКТИКУМ. ОСНОВНЫЕ СВЕДЕНИЯ ОБ ИНФОРМАЦИОННЫХ СИСТЕМАХ (ИС). СИСТЕМЫ БД. МОДЕЛИ ДАННЫХ.

Лекция 1

Белов Александр Владимирович

Содержание курса

1. Основные сведения об информационных системах (ИС).
2. Системы баз данных. Базы данных (БД). Системы управления базы данных (СУБД).
3. Модели данных.
4. Модель «Сущность-Отношение» (E-R).

Информационные системы. Основные понятия

Информационная система (ИУС) – это программно-аппаратный комплекс, предназначенный для обеспечения целенаправленной деятельности конечных пользователей, обеспечивающий сбор, хранение и обработку информации.

Основные функции ИС:

1. Сбор и хранение больших объемов информации;
2. Обработка информации в соответствии с требованиями пользователей
3. Представление данных в виде, удобном для их анализа и принятия управленческих решений на его основе.

Состав ИС:

- База данных как совокупность определенным образом организованных данных
- Программные приложения, реализующие функции системы
- Пользовательский интерфейс

Классификация информационных систем

Тип информационной системы зависит от того, чьи интересы она обслуживает и на каком уровне управления.

Информационные системы можно классифицировать по различным признакам:

- **Масштаб системы** – однопользовательская, групповая (APM), корпоративная
- **Организация (архитектура) системы** – локальные, распределенные
- **Характер задач обработки данных и степень сложности алгоритмов** решения этих задач – OLTP, OLAP, DSS
- **Уровень управления** (характер управленческих решений)
- **Характер обработки и использования информации**
- **Функциональный признак;**
- **Отрасли производства, в которых применяются ИС;**
- **Характер представления и логическая организация данных;**

Уровень управления

- Информационные системы уровня оперативного (операционного) управления – бухгалтерская система, операционный день банка, системы обработки заказов, регистрации билетов;**
- Информационные системы уровня тактического управления (среднее звено) – функции мониторинга, администрирования, контроля, принятия решений (АСУТП, MES-системы и т.п.);**
- Информационные системы стратегического уровня – стратегическое планирование, принятие решений уровня топ-менеджмента (модули крупных ERP-систем, IBM Watson, IBMi2 и т. п.)**

Характер обработки и использования информации

- **Информационно-поисковые системы** производят ввод, систематизацию, хранение, выдачу информации по запросу пользователя без сложных преобразований данных (информационно-поисковая система в библиотеке, в железнодорожных и авиакассах).
- **Управляющие информационные системы** вырабатывают информацию, на основании которой человек принимает решение. Для этих систем характерен тип задач расчетного характера и обработка больших объемов данных. Примером могут служить система оперативного планирования выпуска продукции, система финансового управления компанией.
- **Рекомендательные информационные системы** вырабатывают информацию, которая принимается человеком к сведению и не превращается немедленно в серию конкретных действий. Эти системы обладают более высокой степенью интеллекта, так как для них характерна обработка знаний, а не данных. Примеры таких систем – экспертные системы, системы поддержки принятия решений
- **Гибридные системы** – включающие в себя элементы всех трех типов

Функциональный признак



Отраслевой признак

- Автоматизированные банковские системы;
 - АИС для страховых компаний;
 - Системы для торговли;
 - Системы для предприятий машиностроения;
 - Системы для предприятий нефтегазовой отрасли;
-
-

Основные определения

Определение 1

База данных (БД) – это система специальным образом организованных данных в соответствии с концептуальной схемой, описывающей характеристики этих данных и взаимоотношения между ними.

Определение 2

Система управления базой данных (СУБД) – совокупность специальных языковых и программных средств, позволяющих выполнять операции, связанные с хранением, доступом и корректированием информации.

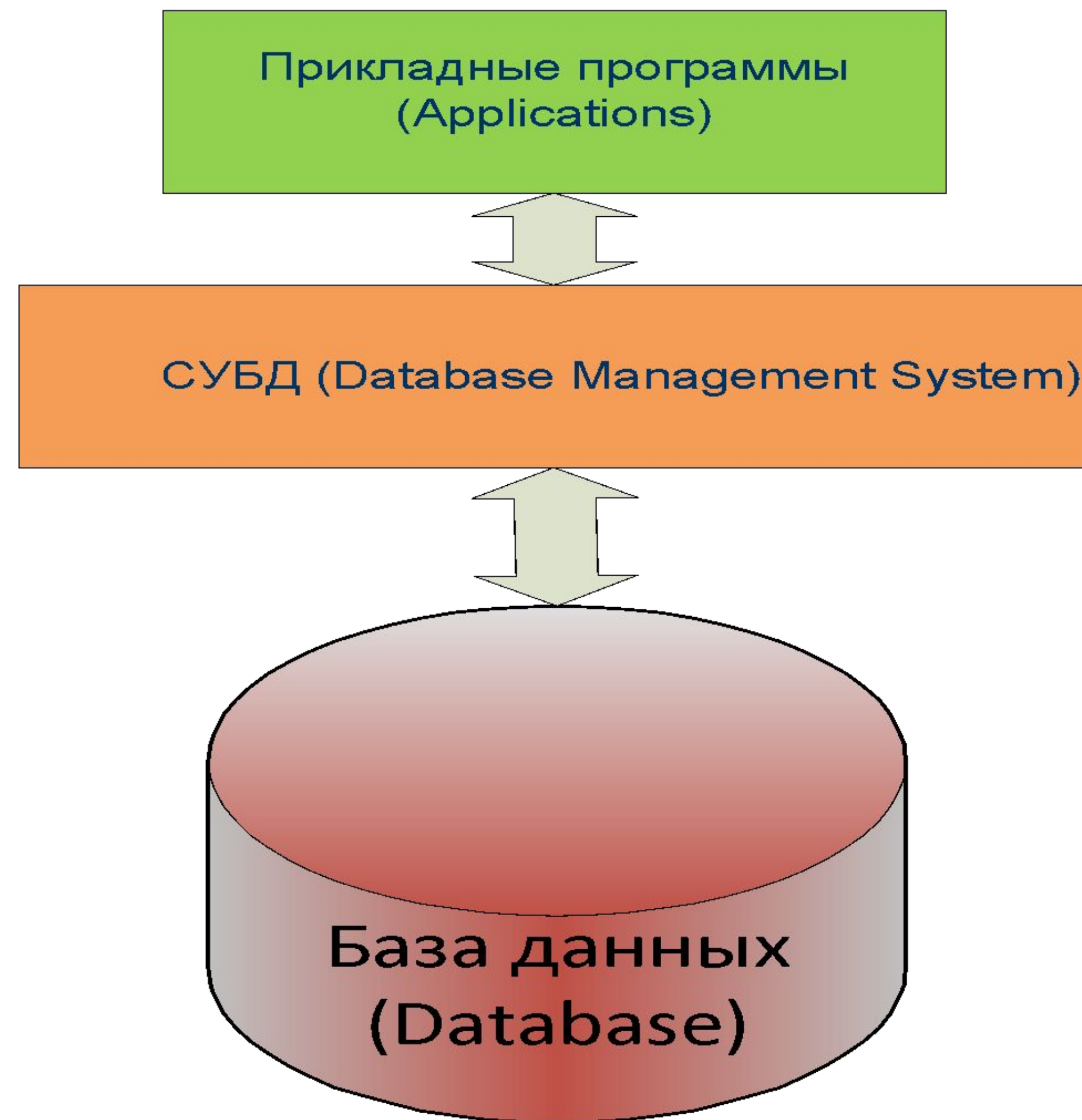
Система баз данных (Database Systems)

Система БД – это система специальным образом организованных данных, программных, технических, языковых, организационно-технических средств, предназначенных для обеспечения централизованного накопления и коллективного многоцелевого использования данных.

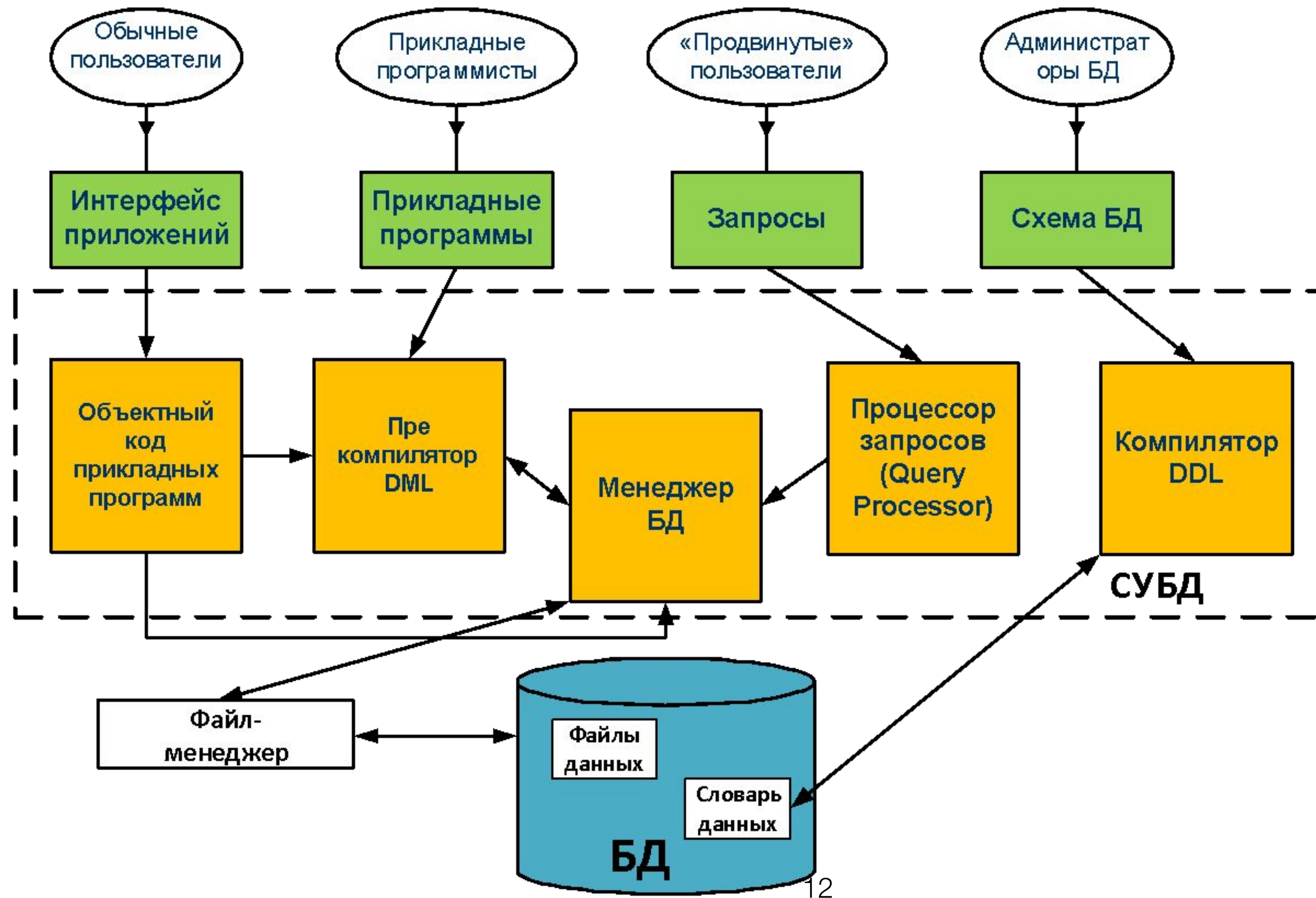
Преимущества использования систем БД:

- Интегрированное хранение данных сокращает избыточность хранимых данных, а это приводит к сокращению затрат не только на создание и хранение данных, но и на поддержание их в актуальном состоянии
- Централизованное управление данными освобождает пользователей от администрирования БД и приводит к снижению трудоемкости создания СБД;
- Независимость прикладных программ от данных;
- Наличие средств, ориентированных на различные категории пользователей.

Структура систем баз данных



Структура систем баз данных



Менеджер БД

Менеджер БД – программа, обеспечивающая интерфейс между данными, хранящимися в БД и прикладными программами и/или запросами. Отвечает за:

- ☐ Взаимодействие с файл-менеджером
- ☐ Обеспечение целостности (проверка ограничений целостности)
- ☐ Обеспечение защиты данных (проверка требований безопасности)
- ☐ Создание резервных копий и восстановление БД
- ☐ Обеспечение параллельной работы пользователей

Язык определения данных (DDL)

- DDL – это множество операторов, с помощью которых можно описать схему БД
- Результат компиляции DDL-операторов – множество таблиц, которые хранятся в Словаре данных (Data Dictionary)
- Data Dictionary содержит метаданные – «данные о данных»
- Структура данных (физическая модель) и методы обработки хранимых данных – специальные определения DDL.
- Информация, хранящаяся в DD, скрыта от пользователей и доступна архитекторам, администраторам БД и прикладным программистам

Язык манипулирования данными (DML)

- Data Manipulation Language – язык, с помощью которого пользователи могут обрабатывать и манипулировать данными, хранящимися в БД
- Манипулирование данными – это операции:
 - ✓ выборки данных по указанным критериям;
 - ✓ удаления данных из БД
 - ✓ вставки данных в БД
 - ✓ модификации (редактирование) данных

Язык манипулирования данными часто называют **языком запросов (Query Language)**

Остальные компоненты СБД

- Файл-менеджер – управляет размещением памяти на диске под информацию БД
- Процессор запросов (Query Processor) – переводит инструкции языка запросов в низкоуровневые команды для менеджера БД
- Пре-компилятор DML – переводит команды DML, включенные в приложения, в обычные процедуры языков высокого уровня
- Компилятор DDL – переводит операторы DDL в множество таблиц, содержащих метаданные
- Словарь данных (Data Dictionary) – хранит метаданные о структуре БД

Модели данных

Определение. Под **моделью данных** будем понимать набор средств, описывающих типы данных, отношения между данными, семантику данных и ограничения на данные

- Логические модели, основанные на понятии «объект» (object-based logical models)
- Логические модели, основанные на понятии «запись» (record-based logical models)
- Физические модели

Понятие логической модели данных

Логическая модель данных является визуальным представлением структур данных, их атрибутов и бизнес-правил. Логическая модель представляет данные таким образом, чтобы они легко воспринимались бизнес-пользователями.

Проектирование логической модели должно быть свободно от требований платформы и языка реализации или способа дальнейшего использования данных.

Логическая модель, основанная на понятии «Запись»

- **Сетевая модель**

Сетевая модель данных – это набор записей, описывающих данные, связанных между собой с помощью ссылок (или указателей). Может быть представлена графом.

- **Иерархическая модель**

Иерархическая модель данных – это сетевая модель, где запись представляет собой дерево, т.е. у каждой дочерней записи может быть одна и только одна родительская запись.

- **Реляционная модель**

Реляционная модель данных – это набор таблиц, описывающих данные и отношения между данными

Логическая модель, основанная на понятии «Объект»

- **Модель «сущность-отношение» (E-R model)**

Предметная область описывается с помощью набора сущностей (**entity**), характеризующих основные моделируемые объекты, и отношений (**relationship**) между ними

- **Объектно-ориентированная модель (Object-oriented model)**

Предметная область описывается набором **объектов**. Объект содержит *данные* и *метод*, обрабатывающий эти данные. Объекты, содержащие однотипные данные и методы, могут объединяться в **классы**.

Модель «Сущность-Отношение». Основные понятия

Сущность (Entity) – это объект предметной области, отличающийся от других объектов

Пример. Гражданин, имеющий ИНН, является сущностью, т.к. именно ИНН отличается от всех остальных граждан страны

Атрибут (attribute) – признак, характеризующий сущность

Пример. Сущность «Клиент банка» (Customer) может иметь атрибуты: ИНН, Тип, Название/Ф.И.О., Адрес, Телефон

Домен (domain) – множество разрешенных значений атрибута

Пример. Домен для атрибута «Название/Ф.И.О.» - множество текстовых строк (string)

Множество сущностей (entity set) – это множество экземпляров сущности

БД – это набор множеств сущностей, каждое из которых содержит любое число сущностей одного типа

Модель «Сущность-Отношение». Основные понятия

Отношение (Relationship) – это связь (ассоциация) между сущностями. Связи выражаются глаголами или глагольными фразами, которые описывают эти взаимосвязи.

Множество отношений – это множество экземпляров отношения.

Пример. Рассмотрим два множества сущностей **Customer** и **Account**. Тогда **CustAcct** – это бинарное множество отношений между этими сущностями

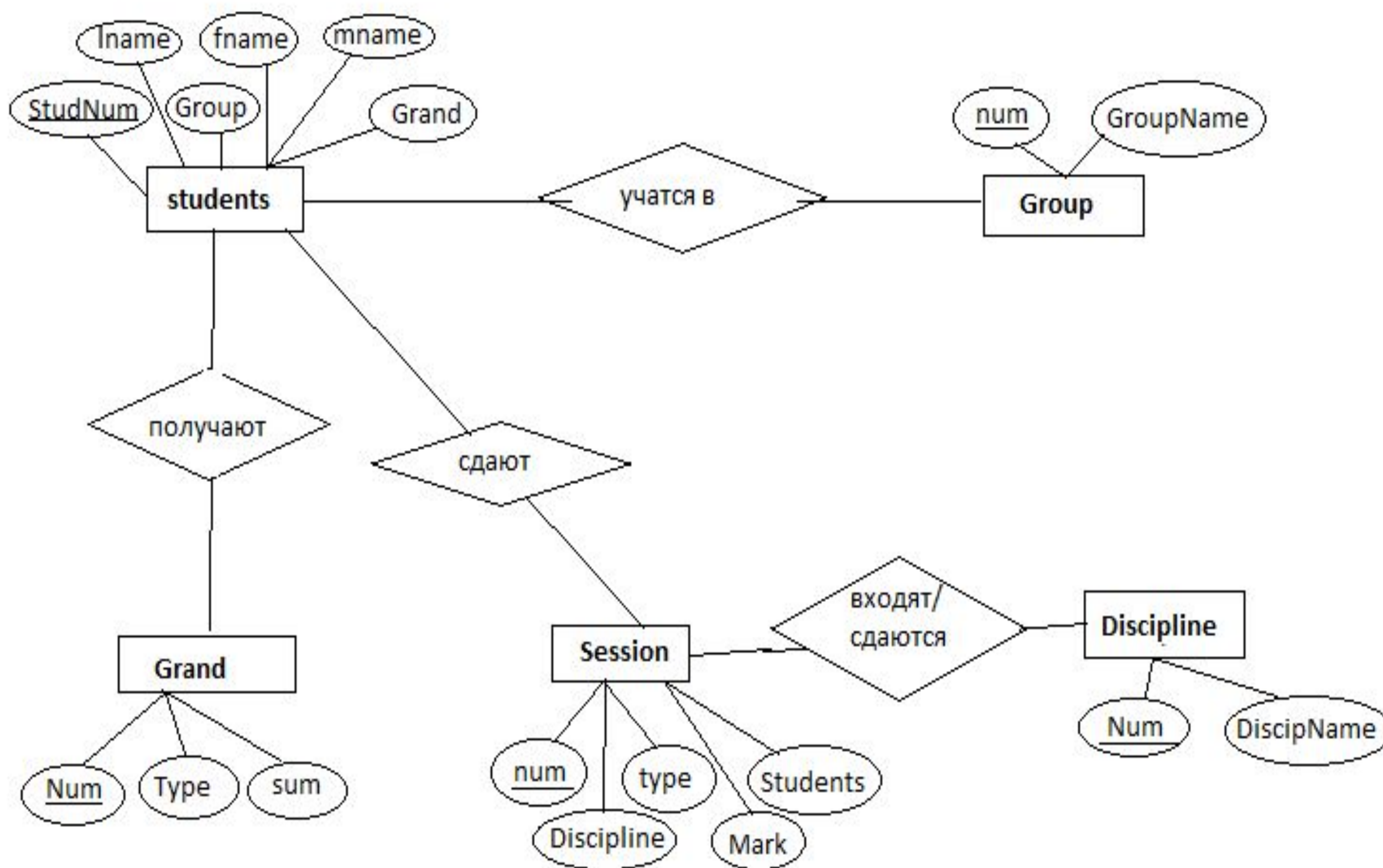
Для бинарного множества отношений **R** между множествами сущностей **A** и **B** **мощностью отношения R** будем называть количество экземпляров сущностей из множества **A**, связанных отношением **R** с экземплярами множества **B**.

Мощность отношений определяется бизнес-правилами (business rules), принятыми в моделируемой предметной области. Виды мощности отношений: «Один к одному» - 1:1, «Один к многим» - 1:N; «Многие к одному» - N:1; «Многие ко многим» - M:N

Пример. Если в банке клиент может иметь несколько счетов, то мощность CustAcct – 1:N

Если в банке существуют корпоративные счета, то мощность CustAcct – M:N

Модель «Сущность-Отношение». E-R - диаграмма



Основные объекты диаграммы

Чена:

Сущность – прямоугольник на диаграмме

Атрибут – эллипс

Отношение – ромб

Все объекты поименованы.

Белов Александр Владимирович

Email: avbelov@hse.ru



НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
УНИВЕРСИТЕТ