



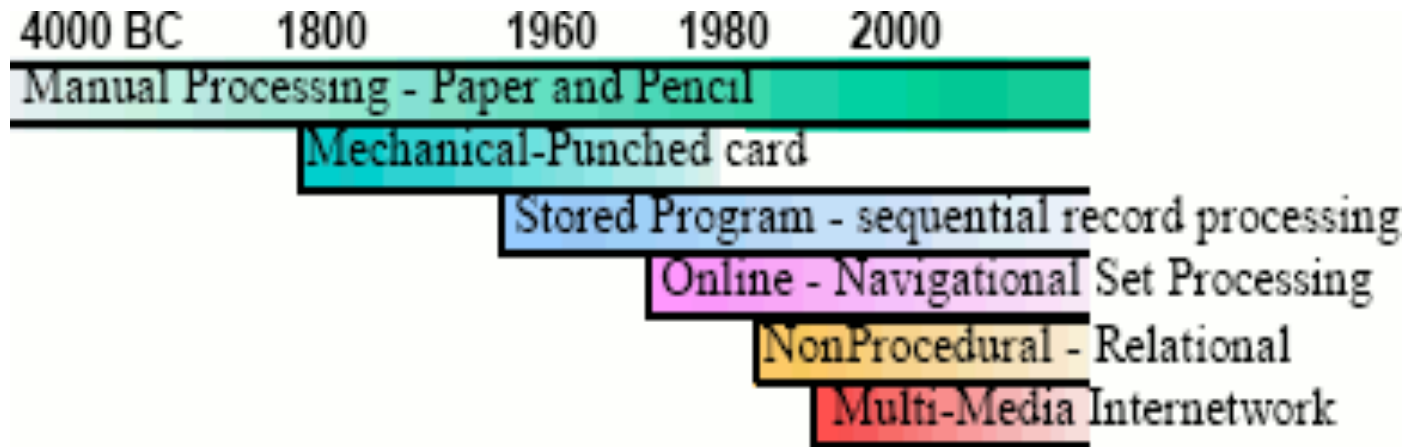
# Базы данных

Введение

# Список литературы

- Карпова И.П. Базы данных. Курс лекций и материалы для практических занятий. – Учебное пособие. – Издательство "Питер", 2013. – 240 с.
- Коннолли Т., Бегг К. Базы данных: проектирование, реализация, сопровождение. Теория и практика, 3-е изд. : Пер. с англ. : Уч. пос. – М.: Изд. дом "Вильямс", 2003. – 1440 с.
- Грабер М. SQL. (любое издание).
- Проектирование реляционных баз данных: Метод. указания к курсовому проектированию по курсу "Базы данных" / МИЭМ; Сост.: Карпова И.П. – М., 2010. – 32 с.
- Изучение основ языка SQL: Метод. указания к лабораторным работам по курсу "Базы данных" / МИЭМ; Сост.: И. П. Карпова. М., 2012. – 32 с.
- Манифест "Системы баз данных третьего поколения". – Журнал «СУБД», ·1995, № 2. – с. 143-159. – URL: <http://rema44.ru/resurs/study/ddb/manifest.html>.
- Манифест «Системы объектно-ориентированных баз данных» // Журнал «СУБД», ·1995, № 4. – с. 142-155. – URL: [http://rema44.ru/resurs/study/ddb/manif\\_oo.html](http://rema44.ru/resurs/study/ddb/manif_oo.html).
- ГОСТ 20886-85. Организация данных в системах обработки данных. Термины и определения.
- ГОСТ 34.320-96. Информационные технологии. Система стандартов по базам данных. Концепции и терминология для концептуальной схемы и информационной базы. – Межгосударственный стандарт. Дата введения 01.07.2001.
- [citforum.ru/database](http://citforum.ru/database)
- [rema44.ru](http://rema44.ru) -> Студентам -> Карпова И.П.

# История



Шесть поколений управления данными, начиная от ручных методов, через несколько стадий автоматизации управления данными.

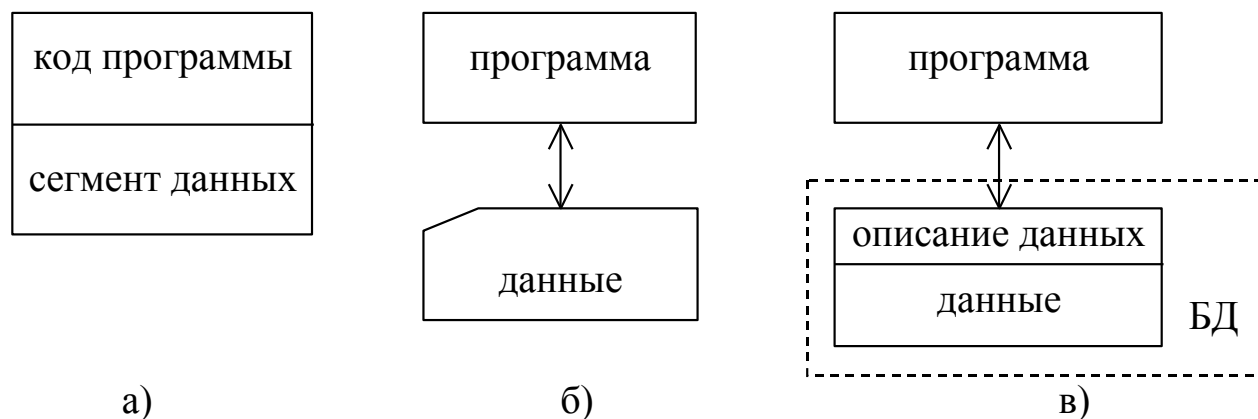
Вначале данные обрабатывались вручную. На следующем шаге использовались оборудование с перфокартами и электромеханические машины для сортировки и табулирования миллионов записей. На третьей фазе данные хранились на магнитных лентах, и сохраняемые программы выполняли пакетную обработку последовательных файлов. На четвертой фазе было введено понятия схемы базы данных и оперативного навигационного доступа к данным. На пятой фазе был обеспечен автоматический доступ к реляционным базам данных и была внедрена распределенная и клиент-серверная обработка. Теперь мы находимся в начале шестого поколения систем, которые хранят более богатые типы данных, в особенности, документы, изображения, аудио- и видеоданные. Эти системы шестого поколения представляют собой базовые средства хранения для появляющихся приложений Internet и intranet.

# Предпосылки появления баз данных

Две основные предпосылки появления баз данных:

- Необходимость хранить и обрабатывать большое количество данных.
- Разработка методов совместного использования данных.

## Эволюция технологий обработки данных:



Массив данных общего пользования в системах, основанных на данных, называется *базой данных*. База данных (БД) является моделью предметной области информационной системы.

# Базы данных: термины

**Информация** – любые сведения о каком-либо событии, объекте или процессе, являющиеся объектом некоторых операций: восприятия, передачи, преобразования, хранения или использования.

**Данные** – это информация, зафиксированная в некоторой форме, пригодной для последующей обработки, передачи и хранения, например, находящаяся в памяти ЭВМ или подготовленная для ввода в ЭВМ.

**Обработка данных** – это совокупность задач, осуществляющих преобразование массивов данных. Обработка данных включает в себя ввод данных в ЭВМ, отбор данных по каким-либо критериям, преобразование структуры данных, перемещение данных на внешней памяти ЭВМ, вывод данных, являющихся результатом решения задач, в табличном или в каком-либо ином удобном для пользователя виде.

**Система обработки данных (СОД)** – это набор аппаратных и программных средств, осуществляющих выполнение задач по управлению данными.

**Управление данными** – совокупность функций обеспечения требуемого представления данных, их накопления и хранения, обновления, удаления, поиска по заданному критерию и выдачи данных. [ГОСТ 20886-85]

# Базы данных: термины

**База данных (БД)** – это совокупность взаимосвязанных структурированных данных, относящихся к определенной предметной области и организованных так, чтобы обеспечить независимость данных от программ обработки.

Фактически база данных – это модель предметной области (ПрО).

**Предметная область (ПрО)** – часть реального мира, подлежащая изучению с целью организации управления и, в конечном итоге, автоматизации.

**Ведение базы данных** – деятельность по обновлению, восстановлению и изменению структуры базы данных с целью обеспечения её целостности, сохранности и эффективности использования [ГОСТ 20886-85].

**Система управления базами данных (СУБД)** – это совокупность программ и языковых средств, предназначенных для управления данными в базе данных, ведения базы данных и обеспечения взаимодействия её с прикладными программами [ГОСТ 20886-85].

**Автоматизированная информационная система (АИС)** представляет собой совокупность данных, экономико-математических методов и моделей, технических, программных средств и специалистов, предназначенную для обработки информации и принятия управленческих решений.

# Банк данных (БНД)

**Банк данных (БНД)** – это автоматизированная информационная система, включающая в свой состав комплекс специальных методов и средств (математических, информационных, программных, языковых, организационных и технических) для поддержания динамической информационной модели предметной области с целью обеспечения информационных запросов пользователей. Банк данных должен:

- Обеспечивать информационные потребности внешних пользователей.
- Обеспечивать возможность хранения и модификации больших объёмов многоаспектных данных.
- Обеспечивать заданный уровень достоверности хранимых данных и их непротиворечивость.
- Обеспечивать доступ к данным только пользователям с соответствующими полномочиями.
- Обеспечивать поиск данных по произвольной группе признаков.
- Удовлетворять заданным требованиям по производительности при обработке запросов.
- Иметь возможность реорганизации при изменении границ ПО.
- Обеспечивать выдачу пользователям данных в различной форме.
- Обеспечивать простоту и удобство обращения внешних пользователей к данным.

# Автоматизированные информационные системы (АИС)

По сферам применения и правилам организации различают **информационно-поисковые (ИПС)** и **системы обработки данных (СОД)**.

В зависимости от характера информационных ресурсов АИС подразделяют на **документальные** и **фактографические** системы.

The screenshot shows a search interface for a library system. The title bar reads "Библиотеки 1.01 - [Поисковая форма]". The menu bar includes "Файл", "Правка", "Вид", "Вставка", "Формат", "Записи", "Сервис", "Окно", and "Справка".

The search form contains the following fields and controls:

- Округ: dropdown menu with "Центральный" selected.
- Номер: empty text input field.
- Найти: button.
- Район: empty dropdown menu.
- Название: text input field containing "\*Горького\*".
- Класс: empty dropdown menu.
- Тип: dropdown menu.
- Система: empty dropdown menu.
- Ведение: empty dropdown menu.
- Очистить: button.
- Reset button (arrow with plus sign).

Below the search form is a table with the following data:

	Id	Номер	Название	Тип
▶	259	43	им. Горького А.М.	
	258	43	им. Горького А.М.	
	248	23	им. Горького А.М.	Детская
*	етчик)			

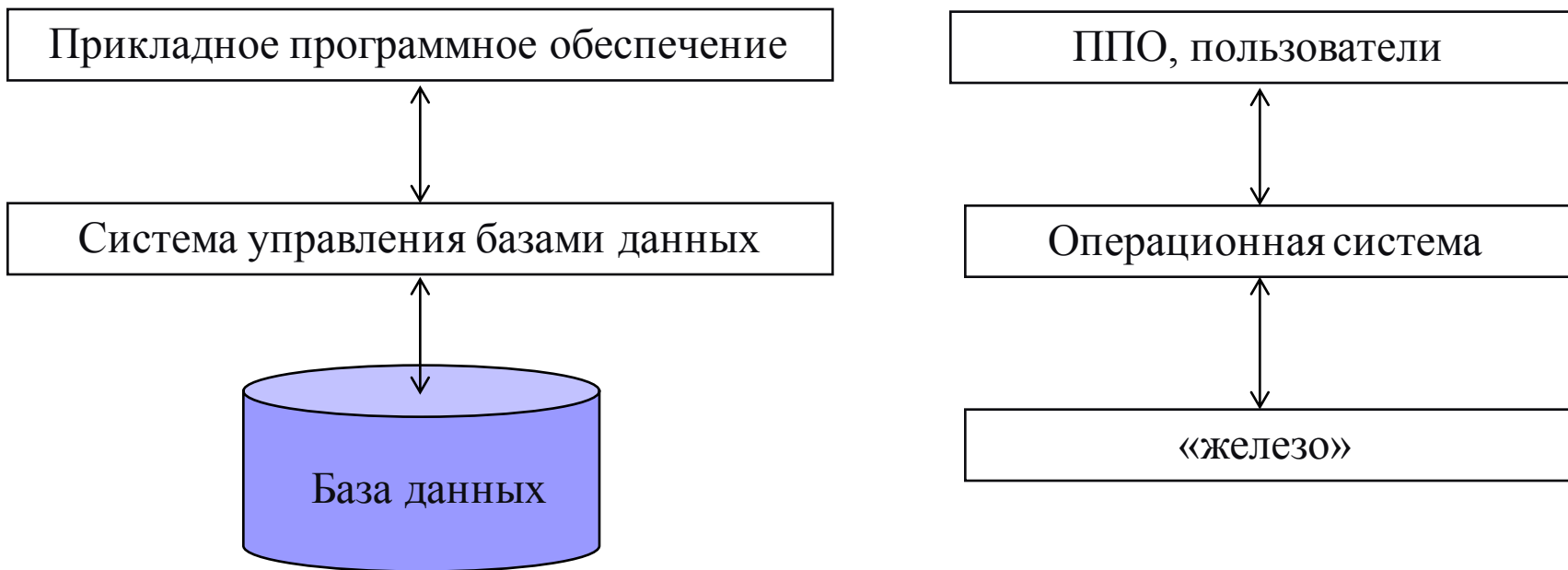
At the bottom, there is a record navigation bar showing "Запись: 1 из 3" with navigation buttons.



# Компоненты системы баз данных

Основным принципом организации базы данных является совместное хранение данных и их описаний. Это отличает базу данных от любого другого набора данных, хранящихся в ЭВМ.

Описание базы данных хранится в так называемом **словаре-справочнике данных (ССД) или каталоге данных**. Хранение данных вместе с их описанием и позволяет обеспечивать независимость данных от программ, т.е. интерпретация данных определяется не программой, а описанием данных.



# Основные функции СУБД

Система БД включает два или три основных компонента:

1. собственно базу данных
2. систему управления базами данных – СУБД
3. программы обработки данных (прикладное программное обеспечение, ППО), которые обращаются к данным через СУБД.

СУБД обеспечивает выполнение двух групп функций:

- предоставление доступа к базе данных прикладному программному обеспечению (или квалифицированным пользователям);
- управление хранением и обработкой данных в БД.

Таким образом, обращение к базе данных возможно только через СУБД.

# Уровни представления данных архитектуры ANSI/SPARC



Концептуальный уровень: поддерживает единый взгляд на базу данных, общий для всех её приложений и независимый от них и от среды хранения.

Внутренний уровень: схема хранения данных в среде хранения.

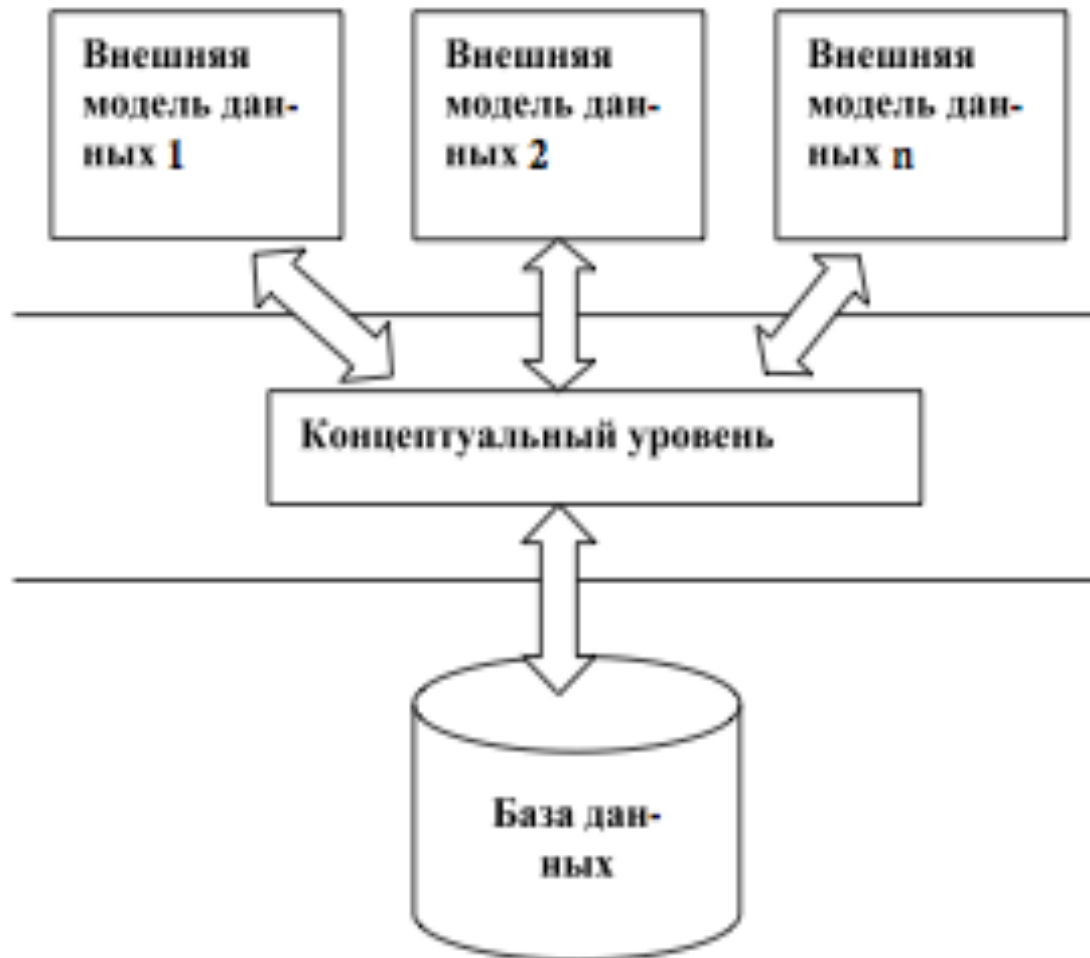
Внешний уровень (внешние схемы): предназначены для групп пользователей.

**Схема базы данных** – это описание базы данных в терминах конкретной модели данных.

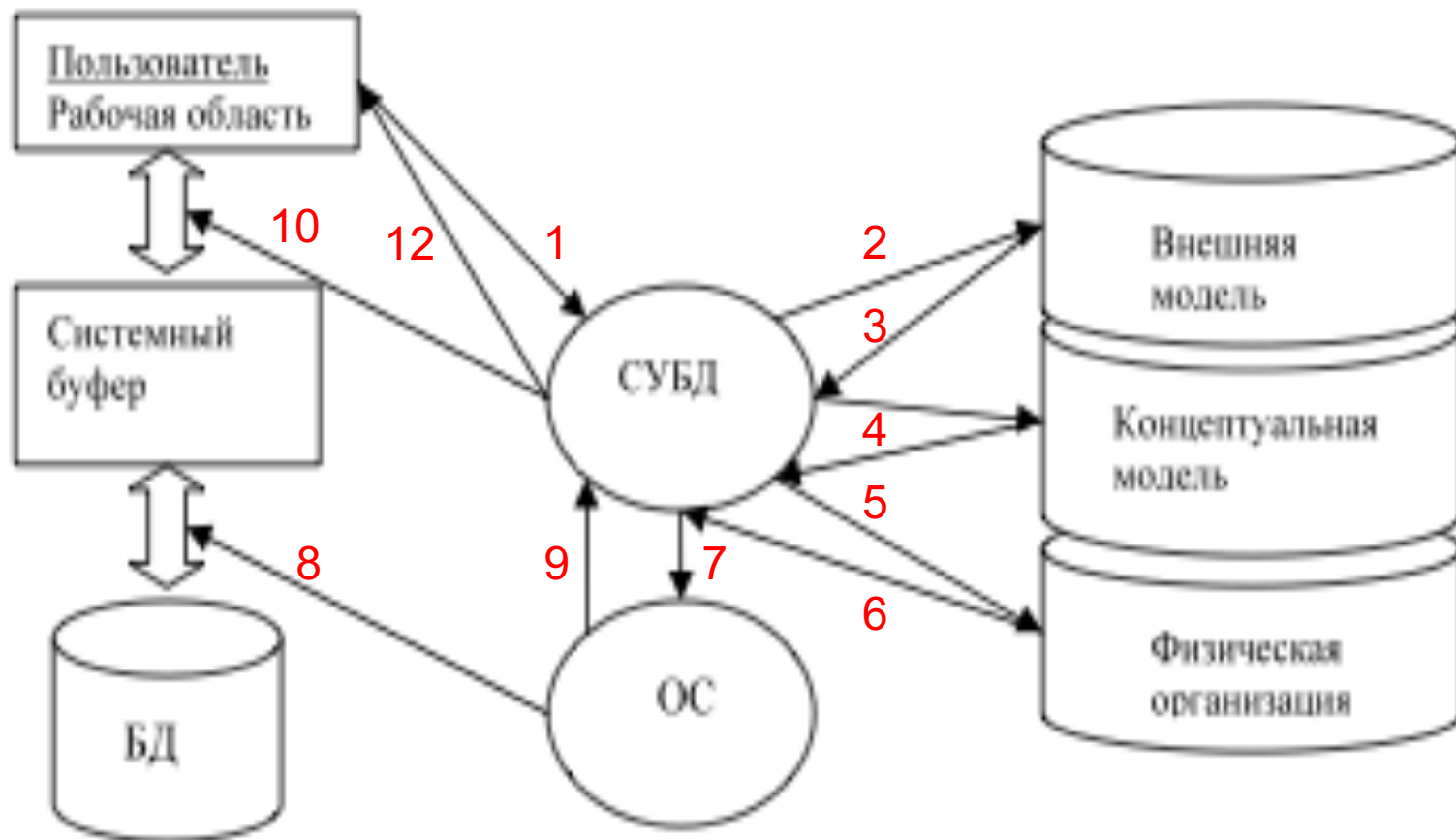
Физическая и логическая независимость данных.


# Трехуровневая модель ANSI/SPARC

## Физическая и логическая независимость



# Прохождение запроса к БД



- 
1. Пользователь посылает СУБД запрос на получение данных из БД.
  2. Анализ прав пользователя и внешней модели данных, соответствующей данному пользователю, подтверждает или запрещает доступ данного пользователя к запрошенным данным.
  3. В случае запрета на доступ к данным СУБД сообщает пользователю об этом (стрелка 12) и прекращает дальнейший процесс обработки данных, в противном случае СУБД определяет часть концептуальной модели, которая затрагивается запросом пользователя.
  4. СУБД получает информацию о запрошенной части концептуальной модели.
  5. СУБД запрашивает информацию о местоположении данных на физическом уровне (файлы или физические адреса).
  6. В СУБД возвращается информация о местоположении данных в терминах операционной системы.
  7. СУБД вежливо просит операционную систему предоставить необходимые данные, используя средства операционной системы.
  8. Операционная система осуществляет перекачку информации из устройств хранения и пересылает ее в системный буфер.
  9. Операционная система оповещает СУБД об окончании пересылки.
  10. СУБД выбирает из доставленной информации, находящейся в системном буфере, только то, что нужно пользователю, и пересылает эти данные в рабочую область пользователя.

# Пользователи банков данных

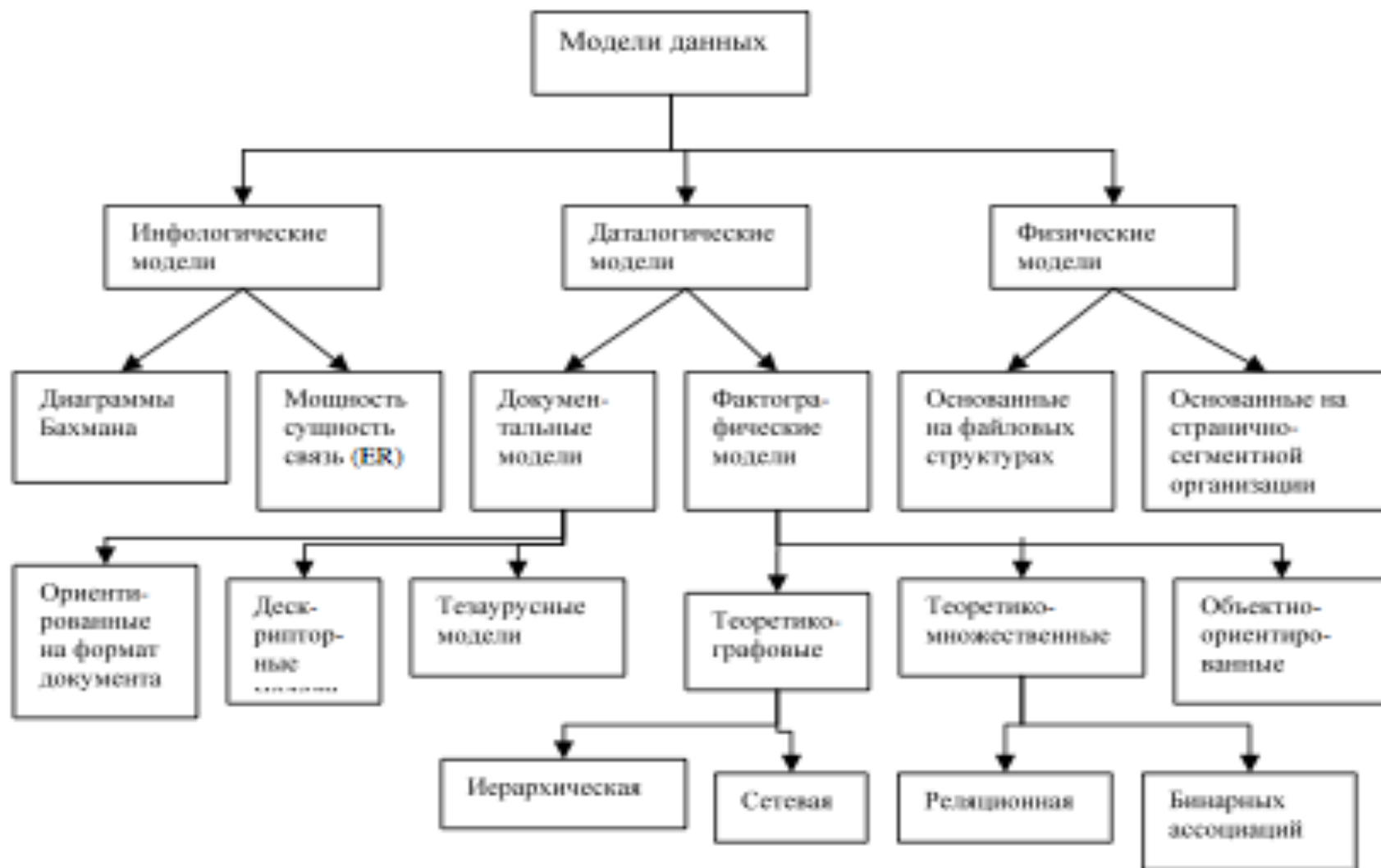
## Жизненный цикл БД

- Проектирование
- Реализация
- Эксплуатация
- Модернизация
- Полная реорганизация

## Виды пользователей

- Конечные пользователи
- Администраторы
- Разработчики

# Классификация моделей данных





# Предметная область. Сущности и атрибуты

Предметная область (ПО) информационной системы рассматривается как совокупность реальных процессов и объектов (**сущностей**), представляющих интерес для её пользователей.

Сущности: базовые и зависимые. Тип сущности. Экземпляр сущности. Для каждого типа сущности необходимо определить имя.

Атрибуты: характеристики сущностей.

Атрибуты бывают:

1. *Идентифицирующие и описательные.*
2. *Составные и простые.*
3. *Однозначные и многозначные.*
4. *Основные и производные.*
5. *Обязательные и необязательные.*

Для каждого атрибута необходимо определить название, указать тип данных и описать ограничения целостности – множество значений, которые может принимать данный атрибут.

# Предметная область. Связи

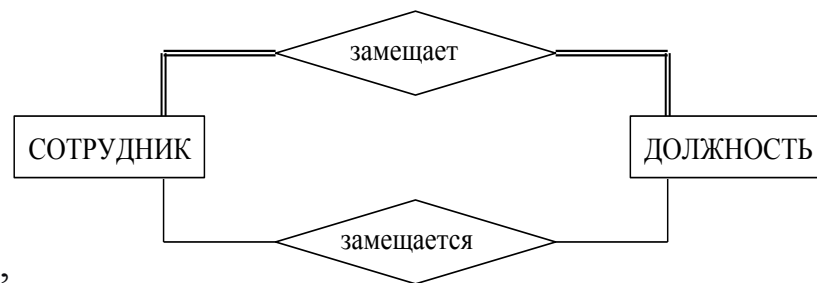
**Связь** – это осмысленная ассоциация между сущностями.

Для связи указывается название, вид (факультативная или обязательная), степень (1:1, 1:n или m:n) и кардинальность (унарная, бинарная, тернарная или n-арная) На практике связи с большими степенями редко используются.

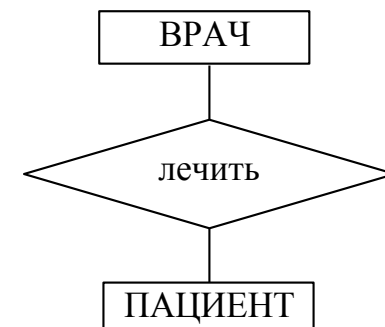
Различают Тип связи и Экземпляр связи.

Тип связи определяется ее именем, обязательностью, степенью и кардинальностью, например, бинарная связь *учится* между сущностями *ГРУППА* и *СТУДЕНТ*, обязательная для студента, кардинальностью 1:n.

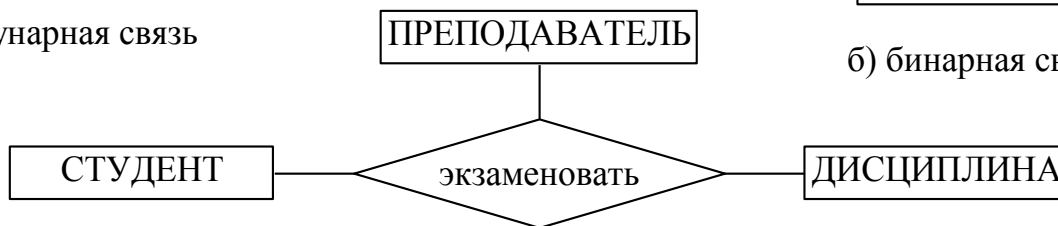
А экземпляр связи – это конкретная связь между студентом Сидоровым и группой Н-11, в которой он учится.



а) унарная связь

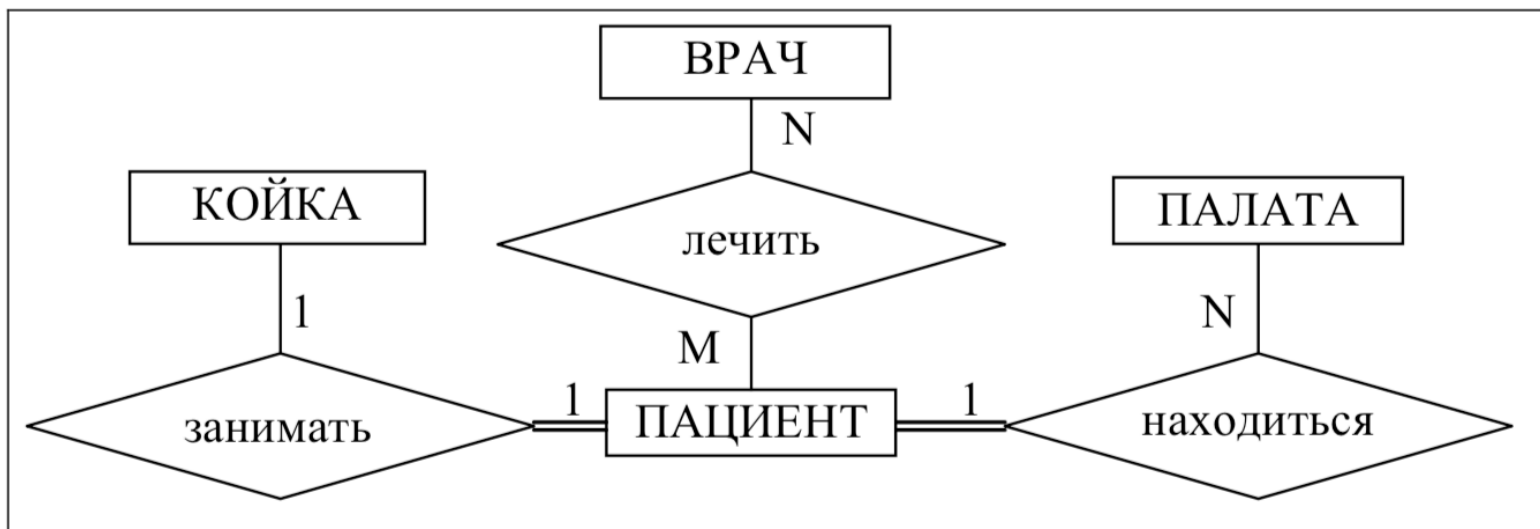


б) бинарная связь



в) тернарная связь

# Предметная область. Связи



Связи, приведённые на рисунке, с учётом семантики означают следующее:

- ✓ пациент–койка (1:1) – каждый пациент занимает одну койку, каждая койка в каждый момент времени может быть занята только одним пациентом;
- ✓ палата–пациент (1:n) – каждый пациент находится в одной палате, в каждой палате могут находиться несколько пациентов;
- ✓ пациент–врач (n:m) – каждый пациент может лечиться у нескольких врачей, каждый врач может лечить несколько пациентов.

# Актуализация данных в БД

Каждому моменту времени можно сопоставить некоторое **состояние предметной области**. Состояния ПО должны подчиняться совокупности правил, которые характеризуют семантику предметной области.

Ограничения целостности.

Для того чтобы обеспечить соответствие базы данных текущему состоянию предметной области, база данных *динамически обновляется* (периодически или в режиме реального времени). Это обновление называется **актуализацией данных**.

Актуализация может проводиться:

- вручную, если изменения в данные вносит пользователь;
- автоматизировано, если изменения инициируются пользователем, но выполняются программно;
- автоматически, если данные поступают в электронном виде и обрабатываются программой без участия человека.

Контроль правильности вносимых изменений.