

Проектирование баз данных

*"Сложная система, спроектированная наспех,
никогда не работает, и исправить её,
чтобы заставить работать,
невозможно".*

Законы Мерфи. 16-й закон системантики



Требования к проекту базы данных

Основные требования, которым должен удовлетворять проект базы данных (БД):

1. Корректность схемы БД.
2. Обеспечение ограничений на ресурсы вычислительной системы.
3. Эффективность функционирования.
4. Обеспечение защиты данных.

5. Гибкость.
6. Простота и удобство эксплуатации.

Удовлетворение первых 4-х требований обязательно для принятия проекта.



Этапы проектирования АИС

В создании АИС (автоматизированной информационной системы) можно выделить следующие этапы:

I. Предпроектная подготовка.

II. Проектирование базы данных.

III. Реализация (создание базы данных и прикладного программного обеспечения, ППО).

I.1. Проектирование начинается обычно с **планирования**, что позволяет:

- разбить задачу на небольшие, независимые, управляемые шаги;
- поставить краткосрочные и долгосрочные цели, которые служат для оценки фактических результатов проектирования и сравнения их с планом;
- определить временные зависимости между задачами, т.е. определить, какие задачи должны быть решены раньше других (составить сетевой план-график работ);
- выявить узкие места, т.е. ресурсы, от которых план зависит сильнее всего;
- спрогнозировать потребности в кадрах для проекта.



Этапы проектирования АИС

Специалисты, необходимые для выполнения этой работы:

- Аналитики (специалисты исследуемой предметной области).
- Пользователи – те работники, для которых создаётся АИС.
- Проектировщики (разработчики базы данных).
- Администраторы (системные, базы данных, безопасности и др.)
- Программисты (разработчики программного обеспечения).

I.2. Определение общих требований к системе подразумевает:

1) Предварительный анализ ПрО.

Включает в себя сбор документов, характеризующих ПрО, укрупнённое описание ПрО (не детализированное) и общую постановку задачи.

В процессе анализа и проектирования желательно ранжировать планируемые функции системы по степени важности. Один из возможных вариантов классификации – MoSCoW-анализ (терминология Клегга и Баркера).



Этапы проектирования АИС

MoSCoW-анализ:

- Must have – необходимые функции;
- Should have – желательные функции;
- Could have – возможные функции;
- Won't have – отсутствующие функции.

2) Рассмотрение и принятие результатов анализа.

3) Определение критических факторов успеха.

В данном случае под термином *критические факторы* подразумеваются как "жизненно важные для приёмки и успешной реализации проекта", так и "критические с точки зрения функционирования системы".

4) Оценка системных ограничений.

В качестве часто встречающихся ограничений можно отметить следующие: финансовые; временные; технические; программные; ограничения, определяемые наличием существующих систем, с которыми необходимо обеспечить совместимость.



Этапы проектирования АИС

5) Определение целевой архитектуры.

6) Определение требований к производительности.

Необходимо примерно оценить количество транзакций в единицу времени и объём обрабатываемых этими транзакциями данных. Требования к производительности зависят от режима, в котором будет функционировать система:

- Интерактивный режим.
- Пакетный режим.
- Режим реального времени.

7) Согласование стандартов проектирования, в частности:

- правил именования объектов;
- стандарта проектной документации;
- правил введения общих типов и т.п.

8) Выбор программных средств для проектирования и реализации системы (имеются в виду вспомогательные средства типа CASE и др.).

I.3. Определение требований пользователей.



Этапы проектирования БД (II)

1. Информационно-логическое (инфологическое) проектирование
 - анализ предметной области;
 - построение модели предметной области;
 - определение границ информационной поддержки;
 - определение групп пользователей.
2. Определение требований к операционной обстановке:
 - выбор аппаратной платформы;
 - выбор операционной системы.
3. Выбор СУБД и других инструментальных программных средств.
 - выбор СУБД;
 - выбор версии СУБД и архитектуры, в которой она будет работать.
4. Логическое проектирование БД (даталогическое):
 - преобразование схемы предметной области в схему базы данных;
 - создание схем отношений;
 - нормализация отношений.
5. Физическое проектирование БД:
 - реализация проекта на DDL-языке выбранной СУБД;
 - создание дополнительных объектов БД (индексов, представлений, триггеров и др.).



Этапы проектирования АИС

III. Реализация.

- 1) Создание прототипа БД и его отладка.
- 2) Разработка и отладка приложений. Выполняется разработчиками программного обеспечения на основе функциональных требований, которые были выявлены на этапах I.2, I.3, и схемы БД.
- 3) Конвертирование и загрузка данных в БД.
- 4) Тестирование работы базы данных и АИС в целом.

Различают такие виды тестов, как:

- **автономные** – тесты отдельных модулей;
- **тесты связей** – тесты между модулями;
- **регрессивные** – тесты на проверку уже протестированных модулей в связи с подключением новых модулей (функций), которые могут нарушить работу ранее созданных модулей;
- **нагрузочные** – тесты на проверку времени реакции системы в рабочем режиме или определение производительности системы;
- **системные** – тесты на проверку функционирования системы в целом;
- **приёмо-сдаточные** – тесты, которые проводятся при сдаче системы (АИС) в эксплуатацию.



Этапы проектирования АИС

III. Реализация.

5) Эксплуатация и сопровождение АИС.

Здесь можно выделить ряд задач:

- В процессе эксплуатации АИС может возникнуть необходимость внесения изменений в систему. Это может быть вызвано изменениями предметной области, появлением новых задач или выявлением существенных недостатков в АИС. Нельзя забывать о том, что все вносимые изменения должны быть документированы.
- Необходимо выполнять резервное копирование данных, чтобы предотвратить их потерю в случае серьёзного сбоя или ошибки пользователя.
- Сопровождение АИС обычно включает периодические проверки выполнения системных ограничений (на объём данных и время реакции системы). В результате этих проверок удаляются устаревшие данные (если не предусмотрено автоматическое архивирование данных). Улучшение показателей производительности системы может быть достигнуто за счёт настройки СУБД, которая выполняется администратором базы данных.



I. Инфологическое проектирование

Инфологическая модель ПрО включает описание структуры и динамики ПрО, характера информационных потребностей пользователей системы. Описание выполняется в терминах, понятных пользователю и независимых от реализации системы.

Обратите внимание: инфологическая модель ПрО не должна зависеть от модели данных, которая будет использована при создании БД.

1. Определение границ предметной области (ПрО).

2. Анализ ПрО.

Выполняется на основе **документов** с помощью **специалистов** в данной ПрО.

3. Методы анализа:

- * функциональный,
- * предметный;
- * метод сущность-связь – entity-relation method, ER-метод.

4. ER-метод, основные понятия:

сущность – объект ПрО, сведения о котором необходимо хранить в БД;

атрибут – характеристика сущности (свойство сущности);

связь – устойчивая ассоциация между сущностями.



Анализ ПрО с помощью ER-метода

Сущности:

✓ **базовые** (наличие базовых сущностей *не зависит* от наличия или отсутствия других сущностей).

✓ **зависимые** (наличие зависимых сущностей *зависит* от наличия или отсутствия других сущностей).

Обычно описание ПО выражается в терминах не отдельных сущностей и связей между ними, а их типов, связанных с ними ограничений целостности и тех процессов, которые приводят к переходу ПО из одного состояния в другое. Такое описание может быть представлено любым способом, допускающим однозначную интерпретацию.

Выделяют понятия *тип сущности* и *экземпляр сущности*.

Тип позволяет выделить из всего множества сущностей ПрО группу сущностей, однородных по структуре и поведению (относительно рамок рассматриваемой ПрО).

Данные в БД представлены экземплярами сущностей.



Анализ ПрО с помощью ER-метода

Атрибуты сущностей:

- ✓ **Идентифицирующие и описательные атрибуты.** Идентифицирующие позволяют отличить один экземпляр сущности от другого; описательные заключают в себе интересующие нас свойства сущности.
- ✓ **Составные и простые атрибуты.** Простой атрибут имеет неделимое значение. Составной атрибут является комбинацией нескольких элементов, возможно, принадлежащих разным типам данных (ФИО, адрес и др.).
- ✓ **Однозначные и многозначные атрибуты** (могут иметь соответственно одно или много значений для каждого экземпляра сущности). Например, дата рождения – это однозначный атрибут, а номер телефона – многозначный.
- ✓ **Основные и производные атрибуты.** Значение основного атрибута не зависит от других атрибутов; значение производного атрибута вычисляется на основе значений других атрибутов. Например, возраст вычисляется на основе даты рождения и текущей даты.
- ✓ **Обязательные и необязательные** (первые должны быть указаны при размещении данных в БД, вторые могут не указываться).

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

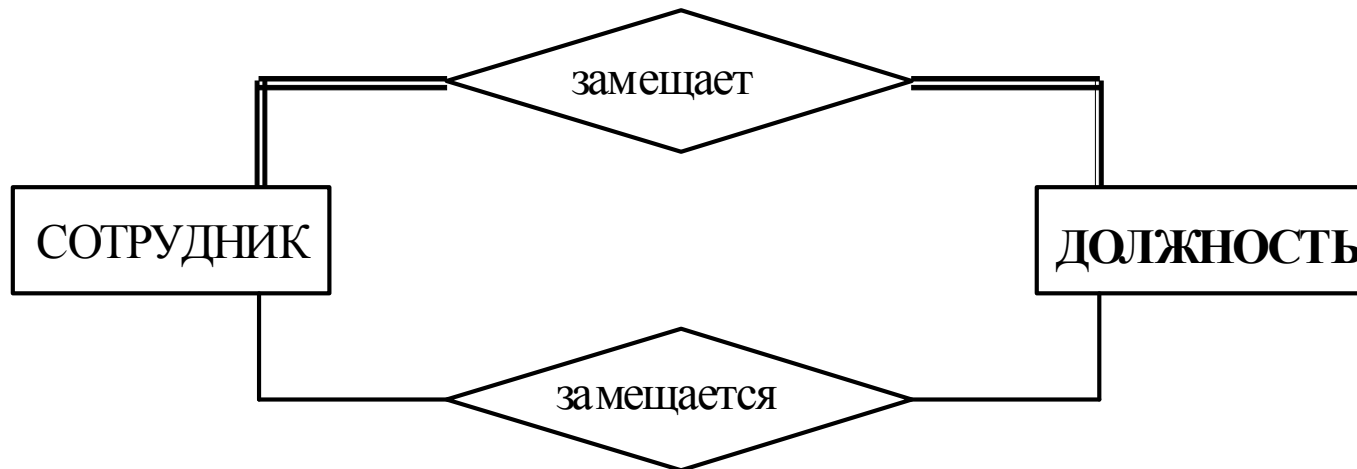
Анализ ПрО с помощью ER-метода

Связи между сущностями:

- ✓ Для связи указывается:
 - название**,
 - тип** (факультативная или обязательная),
 - кардинальность** (1:1, 1:n или m:n),
 - степень** (унарная, бинарная, тернарная или n-арная).

Различают *тип связи* и *экземпляр связи*.

Примеры обязательной и факультативной связей:

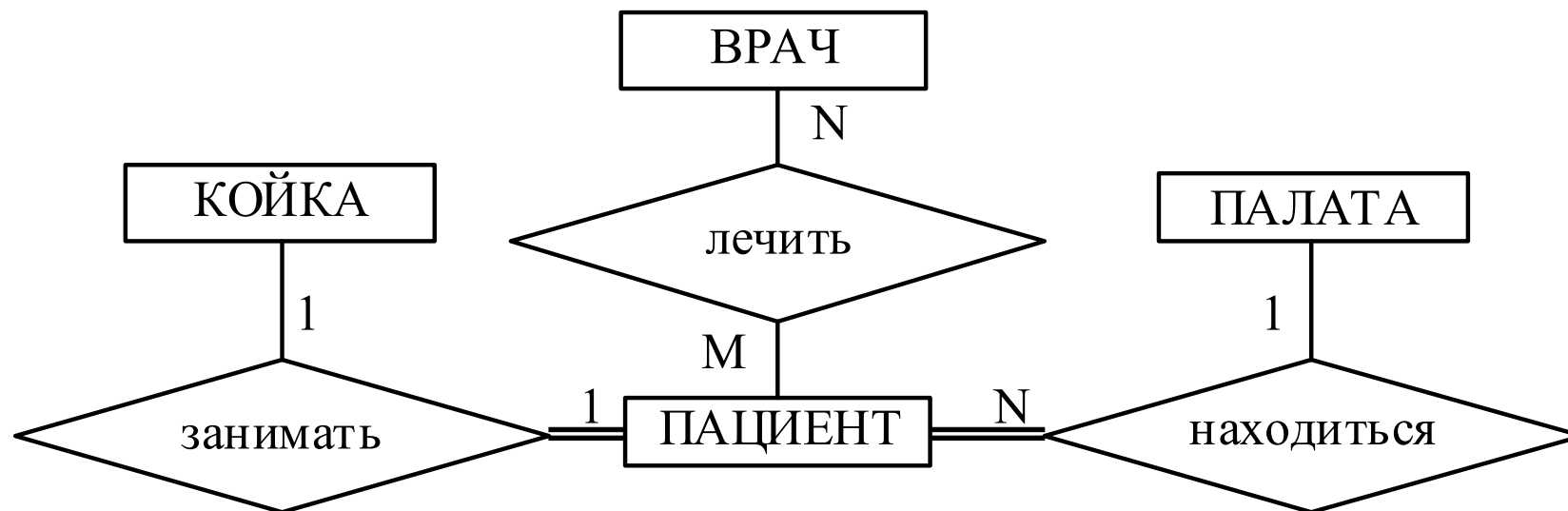


Анализ ПрО с помощью ER-метода

Кардинальность связей между сущностями:

- ✓ один-к-одному (1:1);
- ✓ один-ко-многим (1:n);
- ✓ многие-ко-многим (m:n).

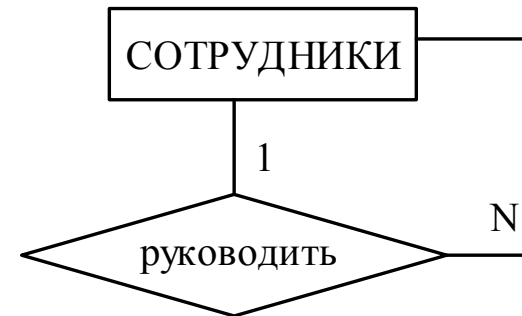
Примеры связей разной кардинальности:



Анализ ПрО с помощью ER-метода

Степень связей между сущностями:

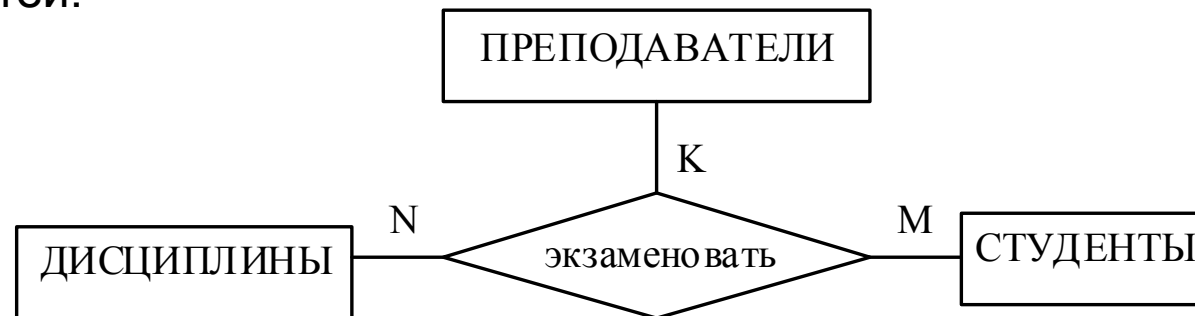
✓ **унарная** – связь между разными экземплярами сущностей одного типа:



✓ **бинарная** – связь между двумя разными типами сущностей:



✓ **тернарная** – связь между тремя разными типами сущностей:





Модель предметной области

Совокупность типов сущностей и типов связей между ними характеризует **структуру предметной области**.

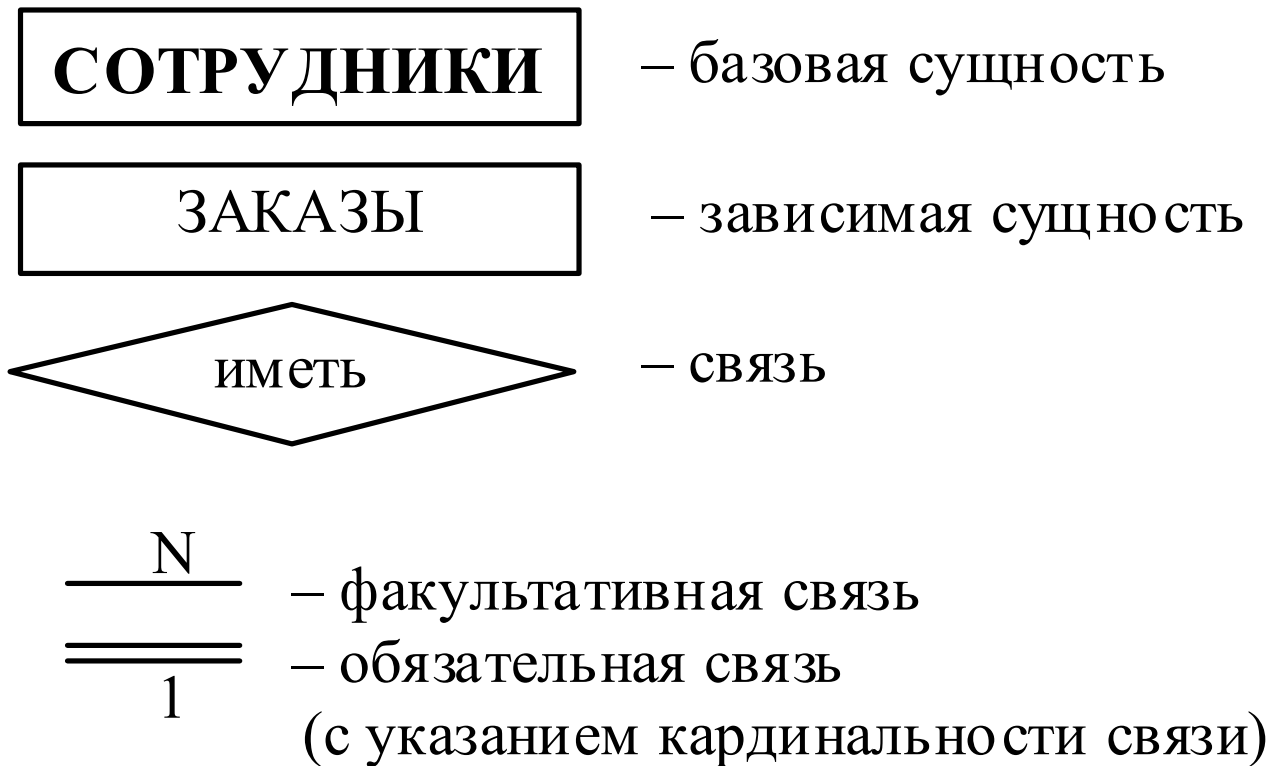
Собственно данные представлены **экземплярами** сущностей и связей между ними. Данные экземпляров сущностей и связей хранятся в базе данных информационной системы, а описание типов сущностей и связей является **метаданными**.

Множества экземпляров сущностей, значения атрибутов сущностей и экземпляры связей между ними могут изменяться во времени. Поэтому каждому моменту времени можно сопоставить некоторое **состояние предметной области**.

Состояния ПО должны подчиняться совокупности правил, которые характеризуют семантику предметной области. В базе данных эти правила могут быть заданы с помощью так называемых *ограничений целостности*, которые накладываются на атрибуты сущностей, типы сущностей, типы связей и/или их экземпляры.

Фактически **ограничения целостности** – это правила, которым должны удовлетворять значения данных в БД.

Обозначения, используемые в ER-диаграммах





Моделирование локальных представлений

Если ПрО содержит много сущностей (10 и более), то она разбивается на ряд локальных областей (**локальных представлений**) по 6-7 сущностей.

Каждое локальное представление включает в себя информацию, достаточную для обеспечения информационных потребностей одной группы будущих пользователей или решения отдельной задачи.

Каждое локальное представление моделируется отдельно, а затем выполняется их объединение (за 1 шаг по попарно).

При объединении локальных представлений используют концепции:

- ✓ **Идентичность.** Два или более элементов модели идентичны, если они имеют одинаковое семантическое значение.
- ✓ **Агрегация.** Позволяет рассматривать связь между элементами как новый элемент.
- ✓ **Обобщение.** Позволяет образовывать многоуровневую иерархию обобщений.

На этапе объединения локальных представлений необходимо устранить все противоречия.

Объединение локальных представлений

Использование обобщения:

Например, пусть в объединяемых представлениях присутствуют следующие сущности:

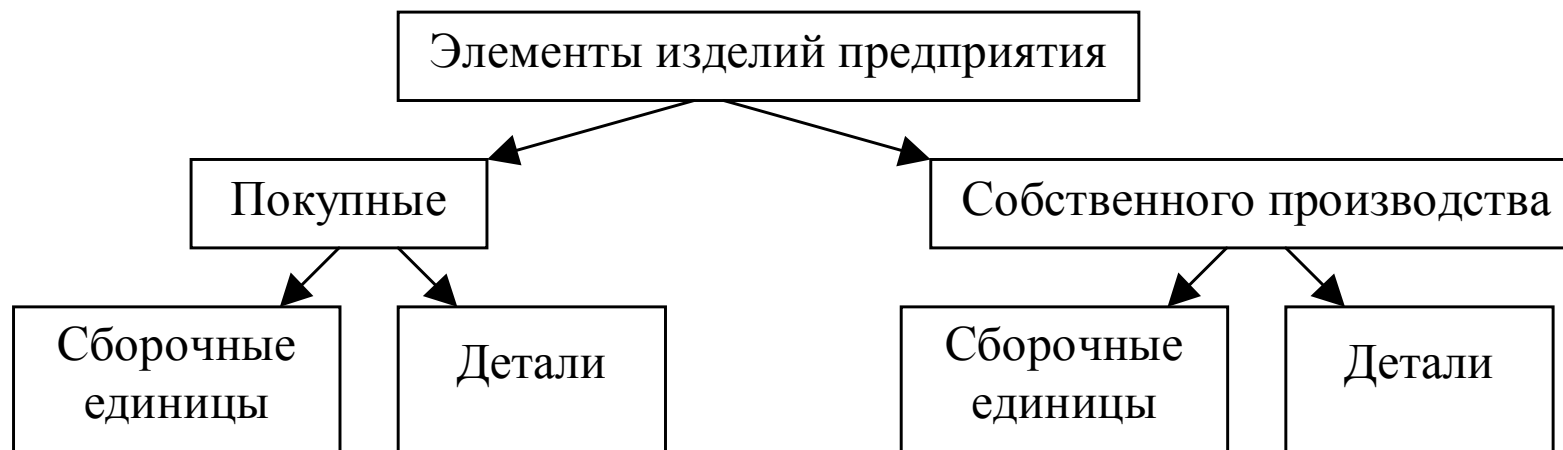
ДЕТАЛИ СОБСТВЕННОГО ПРОИЗВОДСТВА

ДЕТАЛИ ПОКУПНЫЕ

СБОРОЧНЫЕ ЕДИНИЦЫ ПОКУПНЫЕ

СБОРОЧНЫЕ ЕДИНИЦЫ СОБСТВЕННОГО ПРОИЗВОДСТВА

Их можно объединить так :





Результаты инфологического проектирования

- ✓ **Концептуальная инфологическая модель ПрО.** Она фиксируется в виде общей ER-диаграммы предметной области.
- ✓ **Модели локальных представлений** – это внешние инфологические модели (внешние схемы).
- ✓ **Правила (ограничения) целостности**, которым должны удовлетворять сущности ПО, атрибуты сущностей и связи между ними. Часть этих правил реализуется в схеме базы данных, другие – с помощью программного обеспечения.
- ✓ **Перечень групп пользователей системы.** Каждая группа выполняет определённые задачи и обладает разными правами доступа к системе.
- ✓ **Внешние спецификации функций** (процессов), которые будет выполнять АИС.