

Теоретические сведения

Проектирование базы данных – это процесс, который подразумевает использование определённой технологии.

База данных – это, фактически, модель предметной области (ПрО). Значит, для создания БД надо сначала проанализировать ПрО и создать её модель (это называется **инфологическим проектированием**).

Основой для *анализа предметной области* служат документы, которые отражают ПрО, и информация, которую можно получить от специалистов этой предметной области.

Для анализа берутся те документы, которые имеют отношение к решаемой задаче. Изучение документов позволяет выявить объекты (сущности ПрО) и атрибуты сущностей – данные, которые должны храниться в БД.

Из общения со специалистами необходимо извлечь

сведения об особенностях ПрО, которые позволяют установить ограничения целостности, зависимости и связи между объектами (субъектами) предметной области;

алгоритмы обработки данных и какие задачи ставятся перед информационной системой.

При проектировании баз данных используется метод **сущность–связь**, схема ПрО выполняется в виде **ER–диаграммы (entity-relation diagram, диаграмма «сущность–связь»)**.

После создания модели ПрО определяются **требования к операционной обстановке**: какое аппаратное и программное обеспечение необходимо для реализации БД, осуществляется **выбор СУБД**, под управлением которой будет работать создаваемая база данных.

На следующем этапе – этапе **логического проектирования** – **ER-диаграмма** **формальным способом** преобразуется в схему реляционной базы данных (РБД). На основании схемы РБД и описания сущностей ПрО составляются отношения (таблицы) базы данных.

Потом выполняется **нормализация отношений**. Это необходимо сделать для того, чтобы исключить нарушения логической целостности данных и повысить надёжность и достоверность данных.

В результате всех этих операций создаётся концептуальная схема БД – основной документ для базы данных.

Далее, на этапе **физического проектирования** полученные отношения описываются **на языке DDL (Data definition language)** – языке определения данных, который поддерживается выбранной СУБД.

Также необходимо определить способы хранения данных (кластеризация, хеширование) и способы доступа к данным (индексирование) и создать соответствующие индексы и кластеры (если нужно).

Если пользователей БД можно разделить на группы по характеру решаемых задач, то для каждой группы создаётся свой набор прав доступа к объектам БД.

Последовательность проектирования базы данных

Инфологическое проектирование:

1. **Определение задач, стоящих перед базой данных.**
2. **Сбор и анализ документов, относящихся к исследуемой предметной области.**
3. **Описание особенностей ПрО, которые позволяют установить зависимости и связи между объектами (субъектами) предметной области.**
4. **Создание модели предметной области.**
5. **Определение групп пользователей и перечня задач, стоящих перед каждой группой.**

Требования:

6. **Выбор аппаратной и программной платформы для реализации БД.**

На этом этапе производится оценка требований к вычислительным ресурсам, необходимым для функционирования системы, определение типа и конфигурации конкретной ЭВМ, выбор типа и версии операционной системы.

Выбор системы управления базой данных (СУБД) и других инструментальных программных средств

7. **Выбор СУБД (системы управления базой данных).**

Выбор СУБД осуществляется на основании таких критериев, как тип модели данных и её адекватность потребностям рассматриваемой ПрО; характеристики производительности; набор функциональных возможностей; удобство и надежность СУБД в эксплуатации; стоимость СУБД и дополнительного программного обеспечения.

Логическое проектирование БД

8. **Создание логической схемы БД.**
9. **Создание схем отношений, определение типов данных атрибутов и ограничений целостности.**
10. **Нормализация отношений (до третьей нормальной формы).**
11. **Определение прав доступа пользователей к объектам БД.**

На этапе логического проектирования разрабатывается логическая (концептуальная) структура БД. Для реляционной модели существуют формальные правила, которые позволяют преобразовать инфологическую модель ПрО в виде ER-диаграммы в логическую схему базы данных. Кроме получения схемы БД в целом на этом этапе выполняют создание схем отношений и их нормализацию.

Физическое проектирование БД

12. **Написание текста создания основных объектов базы данных на языке SQL в синтаксисе выбранной СУБД (пользователи, таблицы и др.).**
13. **Написание текста создания вспомогательных объектов базы данных (представления, индексы, триггеры, роли и т.д.). Инфологическое проектирование**

Этап физического проектирования заключается в определении схемы хранения, т.е. физической структуры БД. Схема хранения зависит от той физической структуры, которую поддерживает выбранная СУБД. Физическая структура БД, с одной стороны, должна адекватно отражать логическую структуру БД, а с другой стороны, должна обеспечивать эффективное размещение данных и быстрый доступ к ним. Результаты этого этапа документируются в форме схемы хранения на языке определения данных (DDL, Data Definition Language) выбранной СУБД. Принятые на этом этапе решения оказывают огромное влияние на производительность системы.

Одной из важнейших составляющих проекта базы данных является разработка средств защиты БД. Защита данных имеет два аспекта: защита от сбоев и защита от несанкционированного доступа. Для защиты от сбоев на этапе физического проектирования разрабатывается стратегия резервного копирования. Для защиты от несанкционированного доступа каждому пользователю доступ к данным предоставляется только в соответствии с его правами доступа, набор которых также является составной частью проекта БД.

Инфологического проектирования

Основными задачами этапа инфологического проектирования являются определение предметной области системы и формирование взгляда на неё с позиций сообщества будущих пользователей БД, т.е. информационно-логической модели ПрО.

Инфологическая модель ПрО представляет собой описание структуры и динамики ПрО, характера информационных потребностей пользователей в терминах, понятных пользователю и не зависимых от реализации БД.

Метод "сущность-связь" (entity–relation, ER–method)

Сущность – это объект, о котором в системе будут накапливаться данные.

Для сущности указывается название и тип (сильная или слабая).

Сильные сущности существуют сами по себе, а существование слабых сущностей зависит от существования сильных.

Атрибут – свойство сущности.

Различают:

Идентифицирующие и описательные атрибуты. Идентифицирующие позволяют отличить один экземпляр сущности от другого. Описательные атрибуты включают в себя 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.). 3)

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.

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

Связь – это осмысленная ассоциация между сущностями. Для связи указывается название, тип (факультативная или обязательная), кардинальность (1:1, 1:n или m:n) и степень (унарная, бинарная, тернарная или n-арная).

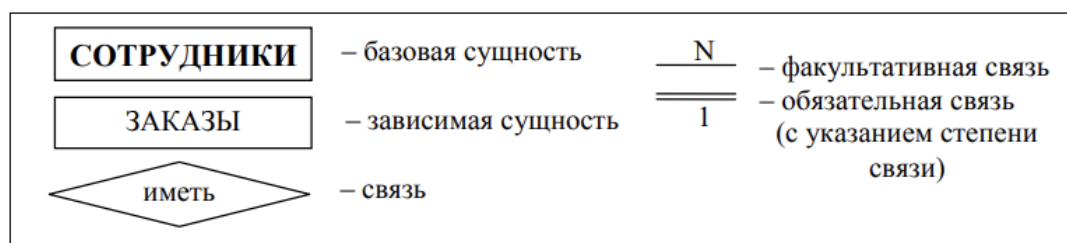


Рис. Обозначения, используемые в ER-диаграммах

Особенности проектирования реляционной базы данных

Проектирование схемы БД должно решать задачи минимизации дублирования данных и упрощения процедур их обработки и обновления. При неправильно спроектированной схеме БД могут возникнуть аномалии модификации данных. Они обусловлены отсутствием средств явного представления типов множественных связей между объектами ПрО и неразвитостью средств описания ограничений целостности на уровне модели данных.

Для решения подобных проблем проводится **нормализация отношений**.

Механизм нормализации реляционных отношений разработал Э.Ф. Кодд (E.F. Codd). Этот механизм позволяет по формальным признакам любое отношение преобразовать к третьей нормальной форме.

Нормализация схемы отношения выполняется путём декомпозиции схемы. **Декомпозицией** схемы отношения R называется замена её совокупностью схем отношений A_i таких, что

$$R = \bigcup A_i,$$

и не требуется, чтобы отношения A_i были непересекающимися.

Первая нормальная форма относится к понятию простого и сложного (составного или многозначного) атрибута.

Первая нормальная форма (1НФ).

Отношение приведено к 1НФ, если все его атрибуты простые.

Для того чтобы привести к 1НФ отношение, содержащее сложные атрибуты, нужно:

разбить составные атрибуты на простые,

построить декартово произведение всех многозначных атрибутов с кортежами, к которым они относятся.

Для идентификации кортежа в этом случае понадобится составной ключ, включающий первичный ключ исходного отношения и все многозначные атрибуты.

Вторая нормальная форма основана на понятии *функциональной зависимости*. Пусть X и Y – атрибуты некоторого отношения. Если в любой момент времени каждому значению X соответствует единственное значение Y, то говорят, что Y функционально зависит от X ($X \rightarrow Y$). Атрибут X в функциональной зависимости ($X \rightarrow Y$) называется *детерминантом* отношения.

В нормализованном отношении ***все неключевые атрибуты функционально зависят от ключа отношения***. Неключевой атрибут функционально полно зависит от составного ключа, если он функционально зависит от ключа, но не находится в функциональной зависимости ни от какой части составного ключа.

Вторая нормальная форма (2НФ).

Отношение находится во 2НФ, если оно приведено к 1НФ и каждый **неключевой атрибут функционально полно зависит от составного первичного ключа**.

(Таким образом, если отношение в 1НФ имеет простой первичный ключ, оно сразу находится во второй нормальной форме). Для того чтобы привести отношение ко 2НФ, нужно:

- построить его проекцию, исключив атрибуты, которые не находятся в функционально полной зависимости от составного первичного ключа;
- построить дополнительно одну или несколько проекций на часть составного ключа и атрибуты, функционально зависящие от этой части ключа.

Третья нормальная форма основана на понятии *транзитивной зависимости*. Пусть X, Y, Z – атрибуты некоторого отношения. При этом $(X \rightarrow Y)$ и $Y \rightarrow Z$, но обратное соответствие отсутствует, т.е. Z не зависит от Y или Y не зависит от X . Тогда говорят, что Z транзитивно зависит от X ($X \twoheadrightarrow Z$).

Третья нормальная форма (3НФ).

Отношение находится в 3НФ, если оно находится во 2НФ и каждый неключевой атрибут нетранзитивно зависит от первичного ключа. Для того чтобы привести отношение к 3НФ, нужно:

- построить проекцию, исключив транзитивно зависящие от ключа атрибуты;
- построить дополнительно одну или несколько проекций на детерминанты исходного отношения и атрибуты, функционально зависящие от них.

Пример:

БД проектной организации.

Основной вид деятельности – выполнение проектов по договорам с заказчиками.

Инфологическое проектирование

1. Анализ предметной области

База данных создаётся для информационного обслуживания руководства организации, руководителей проектов и участников проектов. БД должна содержать данные об отделах организации, сотрудниках и проектах.

В соответствии с предметной областью система строится с учётом следующих особенностей, которые определены нами самостоятельно ниже:

- *Каждый сотрудник работает в определённом отделе, в каждом отделе могут работать несколько сотрудников.*
- *Каждый проект относится к определённому отделу, каждый отдел может отвечать за выполнение нескольких проектов.*
- *Каждый сотрудник может принимать участие в выполнении нескольких проектов, над каждым проектом может трудиться несколько сотрудников.*
- *Для каждого проекта назначается руководитель из числа сотрудников того отдела, к которому относится проект.*
- *Каждый проект должен быть выполнен в заданные сроки, каждый проект может состоять из нескольких этапов. Если проект состоит из одного этапа, то сроки его выполнения должны совпадать со сроками выполнения проекта в целом.*
- *Оклад сотрудника зависит от занимаемой должности, за участие в проектах сотрудник получает дополнительное вознаграждение.*
- *Виды участия сотрудников в проектах: руководитель, консультант, исполнитель.*
- *Каждый отдел занимает одно или несколько помещений (комнат), в каждом помещении может быть один или несколько стационарных телефонов.*
-
- *Примечание. Описания особенностей ПрО должно быть достаточно для того, чтобы создать ER-диаграмму.*

Для создания ER-модели необходимо выделить сущности предметной области и указать их атрибуты:

1) **Отделы.** Атрибуты: название, аббревиатура, комнаты, телефоны.

2) **Сотрудники.** Атрибуты: ФИО, паспортные данные, дата рождения, пол, ИНН (индивидуальный номер налогоплательщика), номер пенсионного страхового свидетельства, адреса, телефоны (рабочий, домашний, мобильный), данные об образовании (вид образования (высшее, среднеспециальное и т.д.), специальность, номер диплома, дата окончания учебного заведения), должность, оклад, *логин (имя пользователя)*.

Примечания: 1. Логин потребуется нам для назначения дифференцированных прав доступа. Так как в задании не предусмотрена полная информационная поддержка сотрудников отдела кадров, поэтому мы не будем отражать в БД такие сведения как дату поступления сотрудника на работу, его переводы с одной должности на другую, уходы в отпуска и т.п.

3) **Проекты.** Атрибуты: номер договора; полное название проекта; сокращённое название проекта; дата подписания договора; заказчик; контактные данные заказчика; дата начала проекта; дата завершения проекта; сумма по проекту; дата реальной сдачи проекта; сумма, полученная по проекту на текущую дату.

4) **Этапы проекта.** Атрибуты: номер по порядку, название, дата начала этапа, дата завершения этапа, форма отчетности, сумма по этапу, дата реальной сдачи этапа; сумма, полученная по этапу на текущую дату.

Исходя из выявленных сущностей, построим ER–диаграмму

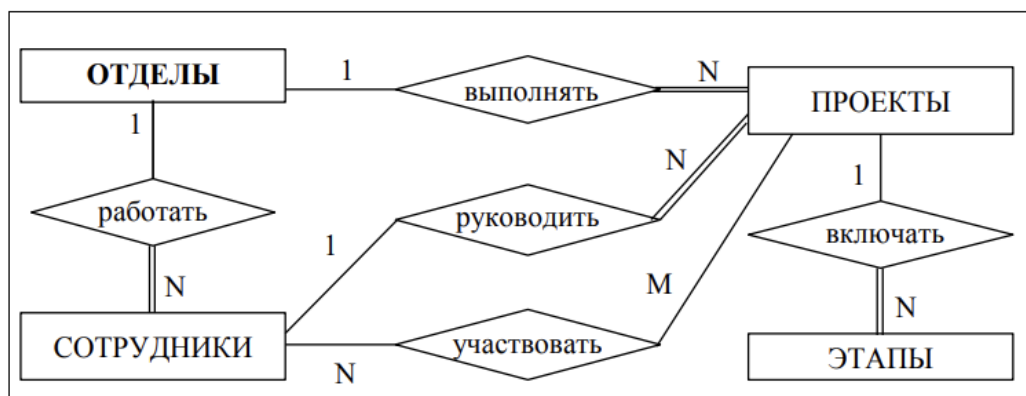


Рис. ER–диаграмма ПрО «Проектная организация»

Анализ информационных задач и круга пользователей системы

Определим группы пользователей, их основные задачи и запросы к БД:

1. Руководители организации:

- заключение новых договоров;
- назначение руководителей проектов;
- получение списка всех участников проектов;
- изменение должностных окладов и штатного расписания;
- получение полной информации о проектах;
- внесение изменений в данные о проектах;
- архивирование данных по завершённым проектам.

Примечание. Архивирование данных в этом пособии подробно не рассматривается. Это сделано для того, чтобы не перегружать схему БД.

2. Руководитель проекта:

- назначение участников проекта;
- получение списка сотрудников, работающих над конкретным проектом; ☐ получение полной информации о проекте, руководителем которого он является; ☐ получение сведений о сотрудниках, которые могут стать участниками проекта;
- определение размера дополнительного вознаграждения сотрудников по конкретному проекту;
- внесение изменений в данные об этапах проекта.

3. Сотрудники отдела кадров:

- приём/увольнение сотрудников; ☐ внесение изменений в данные о сотрудниках.

4. Бухгалтеры:

- получение ведомости на выплату зарплаты.

5. Сотрудники – участники проектов:

- просмотр данных о других участниках проекта;
- просмотр данных о сроках сдачи проекта и форме отчётности.

Определение требований к операционной обстановке

Для выполнения этого этапа необходимо знать (хотя бы ориентировочно) объём работы организации (т.е. количество проектов и сотрудников), а также иметь представление о характере и интенсивности запросов.

Выбор СУБД и других программных средств

Анализ информационных задач показывает, что для реализации требуемых функций подходят почти СУБД MS DQL SERVER

Логическое проектирование реляционной БД

Преобразование ER–диаграммы в схему базы данных

База данных создаётся на основании схемы базы данных. Преобразование ER–диаграммы в схему БД выполняется путем сопоставления каждой сущности и каждой связи, имеющей атрибуты, отношения (таблицы) БД.

Связь типа 1:n (один-ко-многим) между отношениями реализуется через внешний ключ. Ключ вводится для дочернего отношения.

Внешнему ключу должен соответствовать первичный или уникальный ключ основного (родительского) отношения.

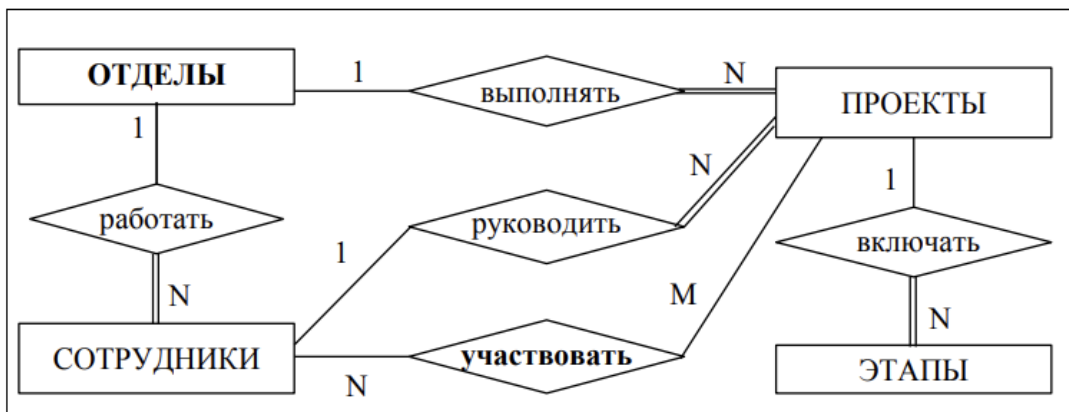


Рис. ER–диаграмма ПрО «Проектная организация»

Связь участвовать между ПРОЕКТАМИ и СОТРУДНИКАМИ принадлежит к типу n:m. Этот тип связи реализуется через вспомогательное отношение Участие, которое содержит комбинации первичных ключей соответствующих исходных отношений.

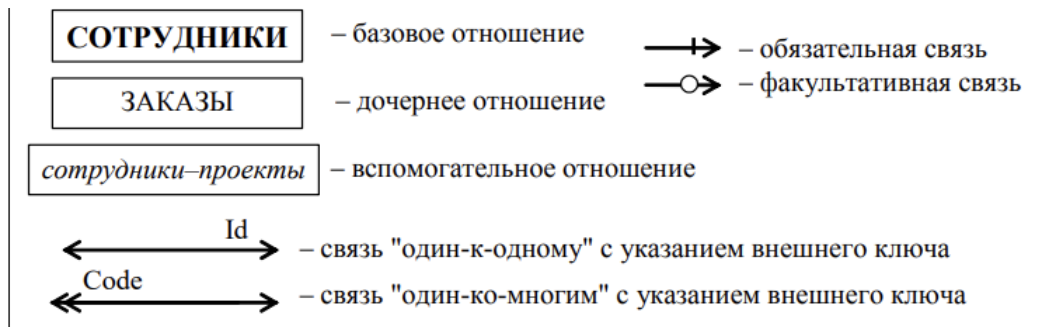


Рис. Обозначения, используемые на схеме базы данных

Клас связи (обязательность). Если сущность одного типа оказывается по необходимости связанной с сущностью другого типа, то между этими типами объектов существует обязательная связь. Иначе связь является факультативной (необязательной)

Полученная схема реляционной базы данных (РБД) приведена на ниже

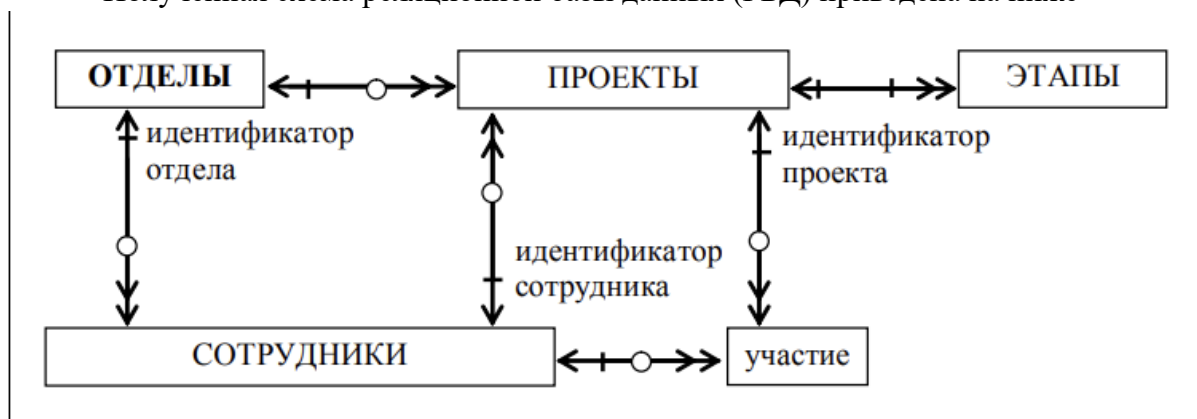


Рис. Схема РБД, полученная из ER–диаграммы проектной организации

Бинарная связь между отношениями не может быть обязательной для обоих отношений. Такой тип связи означает, что, например, прежде чем добавить новый проект в отношение ПРОЕКТЫ, нужно добавить новую строку в отношение ЭТАПЫ, и наоборот. Поэтому для такой связи необходимо снять с одной стороны условие обязательности. Так как все эти связи будут реализованы с помощью внешнего ключа, снимем условие обязательности связей для отношений, содержащих первичные ключи.

Составление реляционных отношений

Каждое реляционное отношение соответствует одной сущности (объекту ПрО) и в него вносятся все атрибуты этой сущности. Для каждого отношения определяются первичный ключ и внешние ключи (в соответствии со схемой БД). В том случае, если базовое отношение не имеет потенциальных ключей, вводится **суррогатный первичный ключ**, который не несёт смысловой нагрузки и служит только для идентификации записей.

Отношения приведены в табл. 1-5. Для каждого отношения указаны атрибуты с их внутренним названием, типом и длиной. Типы данных обозначаются так: N – числовой, С – символьный тип фиксированной длины, V – символьный тип переменной длины, D – дата (этот тип имеет стандартную длину, зависящую от СУБД, поэтому она не указывается). О правилах выбора типов данных подробно рассказано в [1].

Потенциальными ключами отношения ОТДЕЛЫ являются атрибуты Аббревиатура и Название отдела. Первый занимает меньше места, поэтому мы выбираем его в качестве первичного ключа.

Таблица 1. Схема отношения ОТДЕЛЫ (Departs)

Содержание поля	Имя поля	Тип, длина	Примечания
Аббревиатура отдела	D_ID	C(10)	первичный ключ
Название отдела	D_NAME	V(100)	обязательное поле
Комнаты	D_ROOMS	V(20)	обязательное многозначное поле
Телефоны	D_PHONE	V(40)	обязательное многозначное поле

Потенциальными ключами отношения СОТРУДНИКИ являются поля Паспортные данные, ИНН и Номер страхового пенсионного свидетельства. Все они занимают достаточно много места, а паспортные данные кроме того могут меняться. Введём суррогатный первичный ключ Номер сотрудника.

Таблица 2. Схема отношения СОТРУДНИКИ (Employees)

Содержание поля	Имя поля	Тип, длина	Примечания
Номер	E_ID	N(4)	суррогатный первичный ключ
Фамилия, имя, отчество	E_NAME	V(50)	обязательное поле
Дата рождения	E_BORN	D	обязательное поле
Пол	E_GENDER	C(1)	обязательное поле, 'м' или 'ж'
Паспортные данные	E_PASP	V(50)	обязательное поле
ИНН	E_INN	C(12)	обязательное уникальное поле
Номер пенсионного страхового свидетельства	E_PENS	C(14)	обязательное уникальное поле
Отдел	E_DEPART	C(10)	внешний ключ (к Departs)
Должность	E_POST	V(30)	обязательное поле
Оклад	E_SAL	N(8,2)	обязательное поле, > 4500 руб.
Данные об образовании	E_EDU	V(200)	обязательное многозначное поле
Адреса	E_ADDR	V(100)	многозначное поле
Телефоны	E_PHONE	V(30)	многозначное поле
Логин	E_LOGIN	V(30)	

Примечание. Суррогатный первичный ключ также может вводиться в тех случаях, когда потенциальный ключ имеет большой размер (например, длинная символьная строка) или является составным (не менее трёх атрибутов).

В отношении ПРОЕКТЫ три потенциальных ключа: Номер проекта, Название проекта и Сокращённое название. Меньше места занимает первый из них, но он малоинформативен. Зато сокращённое название, используемое в качестве внешнего ключа в других таблицах, позволит специалисту идентифицировать проект без необходимости соединения с отношением ПРОЕКТЫ.

Таблица 3. Схема отношения ПРОЕКТЫ (Projects)

Содержание поля	Имя поля	Тип, длина	Примечания
Номер проекта	P_ID	N(6)	обязательное уникальное поле
Название проекта	P_TITLE	V(100)	обязательное поле
Сокращённое название	P_ABBR	C(10)	первичный ключ
Отдел	P_DEPART	C(10)	внешний ключ (к Departs)
Заказчик	P_COMPANY	V(40)	обязательное поле
Данные заказчика	P_LINKS	V(200)	обязательное поле
Руководитель	P_CHIEF	N(4)	внешний ключ (к Employees)
Дата начала проекта	P_BEGIN	D	обязательное поле
Дата окончания проекта	P_END	D	обязательное поле, больше даты начала проекта
Реальная дата окончания	P_FINISH	D	
Стоимость проекта	P_COST	N(10)	обязательное поле
Полученная сумма	P_SUM	N(10)	обязательное поле, значение по умолчанию – 0

Потенциальным ключом отношения ЭТАПЫ является комбинация внешнего ключа и номера этапа, а потенциальным ключом вспомогательного отношения УЧАСТИЕ является комбинация первых трёх полей этого отношения. Можно вообще не вводить первичный ключ для данных отношений, т.к. на них никто не ссылается. Но уникальность этих комбинации является в данном случае ограничением целостности данных, поэтому мы возьмём эти комбинации в качестве первичных ключей соответствующих отношений.

Таблица 4. Схема отношения ЭТАПЫ ПРОЕКТА (Stages)

Содержание поля	Имя поля	Тип, длина	Примечания	
Проект	S_PRO	C(10)	внешний ключ (к Projects)	составной первичный ключ
Номер этапа	S_NUM	N(2)		
Название этапа	S_TITLE	V(200)	обязательное поле	
Дата начала этапа	S_BEGIN	D	обязательное поле	
Дата окончания этапа	S_END	D	обязательное поле, > даты начала	
Реальная дата окончания	S_FINISH	D	больше даты начала этапа	
Стоимость этапа	S_COST	N(10)	обязательное поле	
Полученная сумма по этапу	S_SUM	N(10)	обязательное поле, значение по умолчанию – 0	
Форма отчётности	S_FORM	V(100)	обязательное поле	

Таблица 5. Схема отношения УЧАСТИЕ (Job)

Содержание поля	Имя поля	Тип, длина	Примечания *	
Проект	J_PRO	C(10)	внешний ключ (к Projects)	составной первичный ключ
Сотрудник	J_EMP	N(4)	внешний ключ (к Employees)	
Роль	J_ROLE	V(20)	'консультант' или 'исполнитель'	
Доплата	J_BONUS	N(2)		

* – в отношении УЧАСТИЕ первичный ключ состоит из первых 3-х полей этого отношения.

Нормализация полученных отношений (до 4НФ)

1НФ. Для приведения таблиц к 1НФ требуется составить прямоугольные таблицы (одно значение атрибута – одна ячейка таблицы) и разбить сложные атрибуты на простые.

Примечание. В реальных БД сложные атрибуты разбиваются на простые, если:

а) этого требует внешнее представление данных;

б) в запросах поиск может осуществляться по отдельной части атрибута.

Разделим атрибут Фамилия, имя, отчество на два атрибута Фамилия и Имя, отчество, Паспортные данные на Номер паспорта (уникальный), Дата выдачи и Кем выдан, Данные об образовании – на Вид образования, Специальность, Номер диплома и Год окончания учебного заведения. Многозначные атрибуты

Комнаты и Телефоны из отношения ОТДЕЛЫ вынесем в отдельное отношение КОМНАТЫ,

а домашние и мобильные телефоны и адреса сотрудников – в отношение АДРЕСА-ТЕЛЕФОНЫ.

Так как в комнате может не быть телефона, *первичный ключ отношения КОМНАТЫ не определен (ПК не может содержать null-значения)*, но на этих атрибутах можно определить составной уникальный ключ. В отношении АДРЕСАТЕЛЕФОНЫ также нет потенциальных ключей: оставим это отношение без первичного ключа, т.к. на это отношение никто не ссылается. Данные об образовании сотрудников также вынесем в отдельное отношение.

Что касается рабочих телефонов сотрудников, то один из этих номеров – основной – определяется рабочим местом сотрудника (мы учитываем только стационарные телефоны). Будем хранить этот номер в атрибуте Рабочий телефон. Наличие других номеров зависит от того, есть ли в том же помещении (комнате) другие сотрудники, имеющие стационарные телефоны. *Добавим в отношение СОТРУДНИКИ атрибут Номер комнаты, чтобы дополнительные номера телефонов сотрудника можно было вычислить из других кортежей с таким же номером комнаты.*

Связь между отношениями СОТРУДНИКИ и КОМНАТЫ реализуем через составной внешний ключ (Номер комнаты, Рабочий телефон).

Мы также удалим вычисляемый атрибут Полученная сумма из отношения ПРОЕКТЫ, т.к. он является суммой значений аналогичного атрибута из отношения ЭТАПЫ ПРОЕКТОВ. Но атрибут Стоимость проекта оставим, т.к. она фигурирует в документации по проекту. А для обеспечения логической целостности данных необходимо предусмотреть в приложении проверку того, что сумма стоимостей по всем этапам совпадает с общей стоимостью проекта.

2НФ. В нашем случае составные первичные ключи имеют отношения ЭТАПЫ ПРОЕКТА и УЧАСТИЕ. Неключевые атрибуты этих отношений функционально полно зависят от составных первичных ключей.

3НФ. В отношении ПРОЕКТЫ атрибут Данные заказчика зависит от атрибута Заказчик, а не от первичного ключа, поэтому его следует вынести в отдельное отношение ЗАКАЗЧИКИ.

Но при этом первичным ключом нового отношения станет атрибут Заказчик, т.е. длинная символьная строка. *Целесообразнее перенести в новое отношение атрибуты Заказчик и Данные заказчика и ввести для него суррогатный ПК.* Так как с каждым заказчиком может быть связано несколько проектов, связь между отношениями ПРОЕКТЫ и ЗАКАЗЧИКИ будет 1:n и суррогатный ПК станет внешним ключом для отношения ПРОЕКТЫ.

В отношении СОТРУДНИКИ атрибут Оклад зависит от атрибута Должность. Поступим с этой транзитивной зависимостью так же, как в предыдущем случае: создадим отношение ДОЛЖНОСТИ, перенесём в него атрибуты Должность и Оклад, а первичным ключом сделаем название должности.

В отношениях СОТРУДНИКИ и ОБРАЗОВАНИЕ атрибуты (Дата выдачи и Кем выдан) и (Номер диплома и Год окончания учебного заведения) зависят не от первичного ключа, а от атрибутов соответственно Номер паспорта и Специальность. Но если мы выделим их в отдельное отношение, то получим связи типа 1:1. Следовательно, здесь декомпозиция нецелесообразна.

4НФ. Отношение АДРЕСА-ТЕЛЕФОНЫ нарушают 4НФ, т.к. не всякий телефон привязан к конкретному адресу (т.е. мы имеем две многозначных зависимости в одном отношении). Но выделять Телефоны в отдельное отношение не стоит, т.к. эти сведения носят справочный характер и не требуется их автоматическая обработка.

Отношения, полученные после нормализации, приведены в табл. 6-15.

Таблица 6. Схема отношения ОТДЕЛЫ (Departs)

Содержание поля	Имя поля	Тип, длина	Примечания
Аббревиатура отдела	D_ID	V(12)	первичный ключ
Название отдела	D_NAME	V(100)	обязательное поле

Таблица 7. Схема отношения КОМНАТЫ (Rooms)

Содержание поля	Имя поля	Тип, длина	Примечания
Отдел	R_DEPART	V(12)	внешний ключ (к Departs)
Номер комнаты	R_ROOM	N(4)	составной уникальный ключ
Телефон	R_PHONE	V(20)	

Таблица 8. Схема отношения ДОЛЖНОСТИ (Posts)

Содержание поля	Имя поля	Тип, длина	Примечания
Название должности	P_POST	V(30)	первичный ключ
Оклад	P_SAL	N(8,2)	обязательное поле, > 4500 руб.

Таблица 9. Схема отношения СОТРУДНИКИ (Employees)

Содержание поля	Имя поля	Тип, длина	Примечания
Идентификатор сотрудника	E_ID	N(4)	суррогатный первичный ключ
Фамилия	E_FNAME	V(25)	обязательное поле
Имя, отчество	E_LNAME	V(30)	обязательное поле
Дата рождения	E_BORN	D	обязательное поле
Пол	E_GENDER	C(1)	обязательное поле
Серия и номер паспорта	E_PASP	C(10)	обязательное уникальное поле
Когда выдан паспорт	E_DATE	D	обязательное поле
Кем выдан паспорт	E_GIVEN	V(50)	обязательное поле
ИНН	E_INN	C(12)	обязательное уникальное поле
Номер пенсионного страхового свидетельства	E_PENS	C(14)	обязательное уникальное поле
Отдел	E_DEPART	V(12)	внешний ключ (к Departs)
Должность	E_POST	V(30)	внешний ключ (к Posts)
Номер комнаты	E_ROOM	N(4)	составной внешний ключ (к Rooms)
Рабочий телефон	E_PHONE	V(20)	
Логин	E_LOGIN	V(30)	

Таблица 10. Схема отношения ОБРАЗОВАНИЕ (Edu)

Содержание поля	Имя поля	Тип, длина	Примечания
Идентификатор сотрудника	U_ID	N(4)	внешний ключ (к Employees)

Вид образования	U_TYPE	V(20)	обязательное поле
Специальность	U_SPEC	V(40)	
Номер диплома	U_DIPLOM	V(15)	
Год окончания учебного заведения	U_YEAR	N(4)	обязательное поле

Таблица 11. Схема отношения АДРЕСА-ТЕЛЕФОНЫ (AdrTel)

Содержание поля	Имя поля	Тип, длина	Примечания
Идентификатор сотрудника	A_ID	N(4)	внешний ключ (к Employees)
Адрес	A_ADDR	V(50)	
Телефон	A_PHONE	V(30)	

Таблицы ОБРАЗОВАНИЕ и АДРЕСА-ТЕЛЕФОНЫ не имеют потенциальных ключей, но мы не будем вводить суррогатные первичные ключи, т.к. на эти таблицы никто не ссылается.

Таблица 12. Схема отношения ЗАКАЗЧИКИ (Clients)

Содержание поля	Имя поля	Тип, длина	Примечания
Номер заказчика	C_ID	N(4)	суррогатный первичный ключ
Заказчик	C_COMPANY	V(40)	обязательное поле
Адрес заказчика	C_ADR	V(50)	обязательное поле
Контактное лицо	C_PERSON	V(50)	обязательное поле
Телефон	C_PHONE	V(30)	

Таблица 13. Схема отношения ПРОЕКТЫ (Projects)

Содержание поля	Имя поля	Тип, длина	Примечания
Номер проекта	P_ID	N(6)	обязательное уникальное поле
Название проекта	P_TITLE	V(100)	обязательное поле
Сокращённое название	P_ABBR	C(10)	первичный ключ
Отдел	P_DEPART	V(12)	внешний ключ (к Departs)
Заказчик	P_COMPANY	N(4)	внешний ключ (к Clients)
Руководитель	P_CHIEF	N(4)	внешний ключ (к Employees)
Дата начала проекта	P_BEGIN	D	обязательное поле
Дата окончания проекта	P_END	D	обязательное поле, больше даты начала проекта
Реальная дата окончания	P_FINISH	D	больше даты начала проекта
Стоимость проекта	P_COST	N(10)	обязательное поле, > 0

Таблица 14. Схема отношения ЭТАПЫ ПРОЕКТА (Stages)

Содержание поля	Имя поля	Тип, длина	Примечания	
Проект	S_PRO	C(10)	внешний ключ (к Projects)	составной первичный ключ
Номер этапа	S_NUM	N(2)		
Название этапа	S_TITLE	V(200)	обязательное поле	
Дата начала этапа	S_BEGIN	D	обязательное поле	
Дата окончания этапа	S_END	D	обязательное поле, больше даты начала этапа	

Реальная дата окончания	S_FINISH	D	больше даты начала этапа
Стоимость этапа	S_COST	N(10)	обязательное поле
Полученная сумма по этапу	S_SUM	N(10)	обязательное поле, значение по умолчанию – 0
Форма отчётности	S_FORM	V(100)	обязательное поле

Таблица 15. Схема отношения УЧАСТИЕ (Job)

Содержание поля	Имя поля	Тип, длина	Примечания	
Проект	J_PRO	C(10)	внешний ключ (к Projects)	составной ПК
Сотрудник	J_EMP	N(4)	внешний ключ (к Employees)	
Роль	J_ROLE	V(20)	обязательное поле	
Доплата	J_BONUS	N(2)		

Схема базы данных после нормализации приведена на рисунке ниже

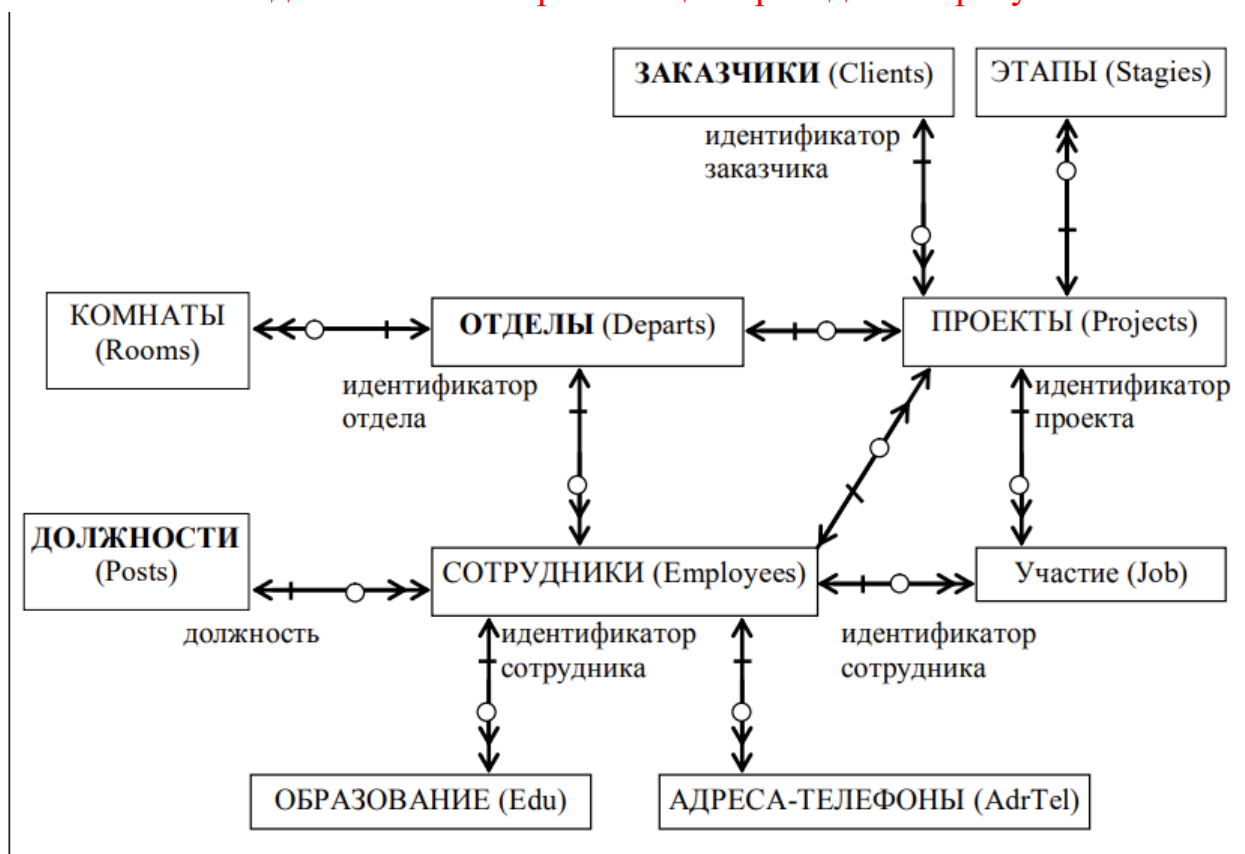


Рис. 6. Окончательная схема БД проектной организации

Определение дополнительных ограничений целостности

Перечислим ограничения целостности, которые не указаны ранее.

1. Атрибут Вид образования может принимать одно из следующих значений: 'начальное', 'среднее', 'средне-специальное', 'высшее'.
2. Атрибут Роль может принимать одно из двух значений: 'исполнитель' или 'консультант'.
3. В поле Доплата хранится величина доплаты сотруднику за участие в проекте (в процентах к его окладу). Значение поля больше либо равно 0.
4. Нумерация в поле Номер этапа начинается с 1 и является непрерывной для каждого проекта.
5. Дата начала первого этапа проекта должна соответствовать началу проекта в целом, дата завершения последнего этапа должна соответствовать завершению проекта в целом. Этапы не должны пересекаться по времени и между ними не должно быть разрывов.
6. Стоимость проекта должна быть равна сумме стоимостей всех этапов этого проекта.