

Лабораторная работа №1

Разработка структуры базы данных

Цель работы: изучение способов задания инфологической модели данных и создания структуры базы данных в заданной предметной области.

Теоретические сведения

Первым этапом разработки структуры базы данных является анализ предметной области. В ходе анализа выявляются основные сущности и их атрибуты, а также взаимосвязи между ними.

Сущность — это любой конкретный или абстрактный объект в рассматриваемой предметной области, информацию о котором необходимо хранить в базе данных.

Атрибут — это именованная характеристика сущности. Наименование атрибута должно быть уникальным для конкретного типа сущности, но может быть одинаковым для различного типа сущностей. Атрибуты используются для определения того, какая информация должна быть собрана о сущности.

Связь — это ассоциирование двух или более сущностей, которое указывает, каким образом связаны сущности. Эта информация необходима для поддержания целостности данных.

На основе проведённого анализа составляется инфологическая модель данных. Одной из форм инфологической модели является диаграмма «сущность — связь», или ER-диаграмма. Основными элементами диаграммы «сущность — связь» являются:

- 1) сущности — представляют собой сущности предметной области. Изображаются прямоугольниками, внутри которых указано наименование сущности;
- 2) ассоциации — представляют собой связи между сущностями. Изображаются ромбами, внутри которых может быть указано наименование связи;
- 3) соединительные линии — соединяют связанные сущности, проходят через вершины ромба, представляющего ассоциацию. Над линиями проставляется степень связи (1 или буква, заменяющая слово «много»).

Между двумя сущностями возможны следующие виды связей: «один-к-одному», «один-ко-многим», «многие-к-одному», «многие-ко-многим». Существуют также более сложные связи, которые встречаются редко и зачастую сводятся к комбинации основных видов связей.

Атрибуты на диаграмму «сущность — связь» допускается не выносить, поскольку изображение большого количества атрибутов увеличивает размер диаграммы, что приводит к уменьшению её наглядности и удобочитаемости. Атрибуты каждой сущности можно указывать в виде отдельного описания, которое можно выполнить в виде приложения к диаграмме.

Для каждой сущности диаграммы «сущность — связь» создаётся таблица базы данных (в некоторых случаях несколько таблиц). Для каждого атрибута сущности создаётся столбец в соответствующей таблице, который также называют полем. Экземпляры сущности хранятся в строках таблицы, которые также называют записями. Каждая таблица должна иметь имя, уникальное для всех объектов базы данных. Имя каждого столбца должно быть уникальным для данной таблицы.

В каждой таблице требуется определить первичный ключ — минимальный набор столбцов таблицы, значения которых однозначно идентифицируют каждую строку этой таблицы. Рекомендуется выбрать в качестве первичного ключа один столбец, соответствующий атрибуту сущности, значения которого являются уникальными для каждой строки таблицы, и требует небольшой объём памяти для их хранения в базе данных. Если в таблице нет подходящих для первичного ключа столбцов, добавляется суррогатный ключ — столбец, не являющийся атрибутом сущности, предназначенный для обеспечения уникальности каждой строки таблицы базы данных. В качестве типа данных обычно выбирается целочисленный тип данных. Значение суррогатного ключа для каждой строки таблицы генерируется автоматически средствами СУБД и называется идентификатором записи.

Для реализации связи между сущностями используется внешний ключ — набор из одного или нескольких столбцов одной таблицы (подчинённой), которая имеет связь с другой таблицей (главной). Количество и типы столбцов внешнего ключа подчинённой таблицы должны совпадать с количеством и типами столбцов первичного ключа главной таблицы.

Если между таблицами установлена связь, то для каждой записи подчинённой таблицы значения полей, соответствующих внешнему ключу, должны совпадать со значениями полей первичного ключа связанной записи главной таблицы. Однако, для любой записи таблицы некоторые её поля могут не содержать значений, если это разрешено в соответствующих столбцах таблицы. В этом случае говорят, что поле содержит пустое значение.

Такая модель была предложена в 1976 году Питером Ченом, им же предложена и самая популярная графическая нотация для модели.

Задания к работе

1. Выполнить анализ предметной области, выделить основные сущности, атрибуты и связи.
2. Создать диаграмму «сущность — связь» в нотации Чена.
3. Самостоятельно изучить нотацию IDEF1X для представления диаграммы «сущность-связь». Создать схему базы данных в нотации IDEF1X.

4. Разработать структуру базы данных и составить описание столбцов таблиц базы данных, включающее: имя столбца, назначение (какие данные хранятся), тип данных, допускает ли столбец пустые значения.

Варианты заданий

Вариант 1. База данных хоккейной лиги. Должна содержать следующие данные: составы команд и информацию о каждом игроке, проведенные игры с информацией о проданных билетах и затраченных средствах. Предусмотреть возможность анализа следующих показателей: популярность команд, рейтинг и эффективность игроков за указанный период.

Вариант 2. База данных сети магазинов продуктов питания. Должна содержать следующие данные: информация о магазинах и имеющихся запасах продуктов, данные о продажах. Предусмотреть возможность анализа следующих показателей: рейтинг популярности товаров, сумма среднего чека по магазину в разное время суток.

Вариант 3. Транспортная компания. База данных должна содержать следующие данные: информацию о водителях, поставщиках товаров и потребителях, завершенные и незавершенные доставки. Предусмотреть возможность анализа следующих показателей: среднее время простоя водителя, соотношение доходности междугородних и местных доставок.

Вариант 4. Сеть книжных магазинов. База данных должна содержать следующие данные: текущие складские запасы печатной продукции, информацию о заказах и продажах. Предусмотреть возможность анализа следующих показателей: наиболее часто заказываемые книги, средний чек по разным группам товаров.

Вариант 5. Сеть автосалонов. База данных должна содержать следующие данные: информацию об автосалонах, продавцах-консультантах, имеющихся в наличии и проданных автомобилях. Предусмотреть возможность анализа следующих показателей: рейтинг продаж для продавцов-консультантов по различным моделям, рекомендации к заказу моделей на основании имеющегося запаса и популярности модели.

Вариант 6. База данных складского комплекса. База данных должна содержать следующие данные: данные о товарах, данные о складах, имеющихся товарах и выполненных и невыполненных заявках на них. Предусмотреть возможность анализа следующих показателей: рейтинг востребованности товаров, рейтинг дефицитности товаров.

Вариант 7. Библиотека. База данных должна содержать следующие данные: информацию об имеющихся изданиях, информацию о читателях,

формуляры для каждого издания. Предусмотреть возможность анализа следующих показателей: выдать рекомендации для читателя с учетом его пола, возраста и прочитанных книг на основании общей статистики.

Вариант 8. Жилищная управляющая компания. База данных должна содержать следующие данные: информацию об исполнителях работ и выполненных работах, жильцах, выставленных им счетах и выполненных ими платежах. Предусмотреть возможность анализа следующих показателей: составить рейтинг злостных неплательщиков, рейтинг исполнителей работ с указанием их доли в статье расходов.

Вариант 9. Гарантийный ремонт. База данных должна содержать следующие данные: информацию о выпускаемых производителем товарах, гарантийных мастерских в разных городах и товарах, ремонт которых они могут производить, данные о выполненных ремонтах. Предусмотреть возможность анализа следующих показателей: рейтинг убыточности товаров, обеспеченность каждого города мастерскими по разным группам товаров и всему ассортименту в целом.

Вариант 10. База данных университета. База данных должна содержать следующие данные: информацию о преподавателях, студентах, их успеваемости, времени и месте проведения занятия. Предусмотреть возможность анализа следующих показателей: количество должников по подразделениям, оценка уровня остепененности преподавателей по подразделениям.

Вариант 11. Веб-форум. База данных должна содержать следующие данные: информацию о зарегистрированных пользователях, создаваемых пользователями темах, данные о сообщениях, которые оставляют пользователи (текст сообщения и автор). Предусмотреть возможность анализа следующих показателей: прирост и отток пользователей, активность пользователей в различных темах за период.

Вариант 12. Страховая компания. База данных должна содержать следующие данные: информацию о клиентах и объектах страхования, виды страховок, информацию о работниках компании и видов страховок, которые они могут предоставлять. Предусмотреть возможность анализа следующих показателей: объем выплат по различным типам страховых случаев за период, рейтинг менеджеров компании в разрезе заключенных контрактов.

Вариант 13. Гостиница. База данных должна содержать следующие данные: информацию о работниках и клиентах, номерах и видах предоставляемых услуг. Предусмотреть возможность анализа следующих показателей: доходность гостиницы в зависимости от сезонности, актуальность предоставляемых услуг.

Вариант 14. Железнодорожные кассы. База данных должна содержать следующие данные: информацию о продажах билетов, посадочных местах, возможных направлениях. Предусмотреть возможность анализа следующих показателей: анализ наиболее популярных направлений в различные времена года, рейтинг поездов.

Вариант 15. Больница. База данных должна содержать следующие данные: информацию о врачах и больных и их диагнозах, видах предоставляемых услуг. Предусмотреть возможность анализа следующих показателей: результаты предоставляемых услуг, статистика пролеченных случаев заболеваний.

Пример выполнения задания

База данных компании, занимающейся перевозкой грузов. Проведем анализ предметной области и выделим следующие сущности:

Заказчик	Лицо, осуществляющее заказ подрядов, одного или нескольких.
Контактные данные	Это информация о том, как связаться с человеком. Она включает в себя ФИО, номер телефона, место фактического пребывания, паспортные данные.
Договор подряда	Каждый подряд зафиксирован документально в виде договора подряда. Это документ, в котором хранится информация такая информация, как статус договора, бюджет подряда, сроки выполнения подряда, агентский процент.
Статус договора	Это состояние, в котором находится договор в данный момент: подписан, отклонен, находится на рассмотрении или нарушен и так далее.
Исполнитель	Это любое лицо, выполняющее работу или оказывающее услуги по заказу, заданию другого лица или согласно договору с заказчиком работ и услуг.

Квалификация	Это соответствие определенным требованиям, таким, как образование, профессиональные знания, навыки и опыт, которые дают возможность специалисту профессионально выполнять работу.
Вид работ	Это принадлежность работы к какому-либо общему виду. К примеру, погрузка груза в вагоны, управление сухогрузом, для морской транспортировки груза.

Выделим следующие связи между сущностями, которые будут храниться в базе данных:

Заказчик заключает **договор подряда** и обладает **контактными данными**.

Договор подряда имеет определенный **статус договора** и информацию о **виде работ**.

Исполнитель имеет **контактные данные**, а также обладает необходимой для данной работы **квалификацией**.

В качестве стандартной графической нотации, с помощью которой можно визуализировать полученную ER-модель, используем диаграмму «сущность-связь». ER-модель в наибольшей степени согласуется с концепцией объектно-ориентированного проектирования, которая в настоящий момент является базовой для разработки сложных программных систем, поэтому многие понятия вам могут показаться знакомыми, и если это действительно так, то тем проще вам будет освоить технологию проектирования баз данных, основанную на ER-модели.

Изобразим диаграмму «сущность-связь» предметной области, опираясь на список сущностей и их атрибутов, описание связей между сущностями, укажем типы отношений:

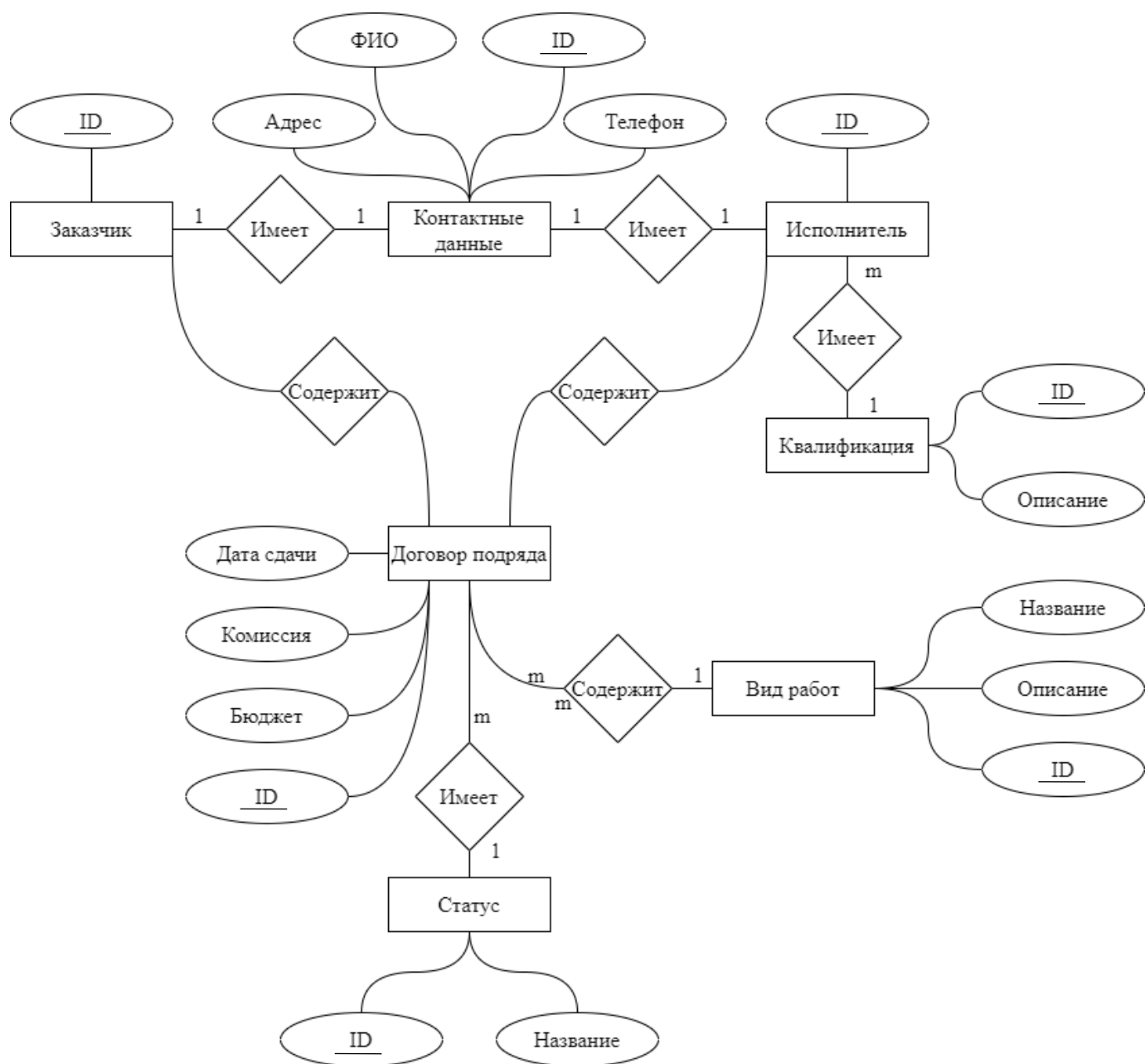


Рис. 1. Диаграмма «сущность-связь» предметной области «Компания, занимающаяся перевозкой грузов»

Для построения инфологической модели базы данных необходимо четко разграничить понятия запроса на данные и ведения данных (ввод, внесение изменений и удаление). Также важно не забывать, что часто база данных является информационной основой для нескольких (а не одного) приложений, часть из которых будет создана в будущем. При плохом проектировании базы данных не представляется возможным исправление такого проекта при помощи каких-либо приложений.

Нотация IDEF1X использует условный синтаксис, специально разработанный для удобного построения концептуальной схемы. Концептуальной схемой мы называем универсальное представление

структуры данных в рамках коммерческого предприятия, независимое от конечной реализации базы данных и аппаратной платформы. При построении данной схемы важно обозначать тип связей между сущностями.

Схема базы данных в нотации IDEF1X:

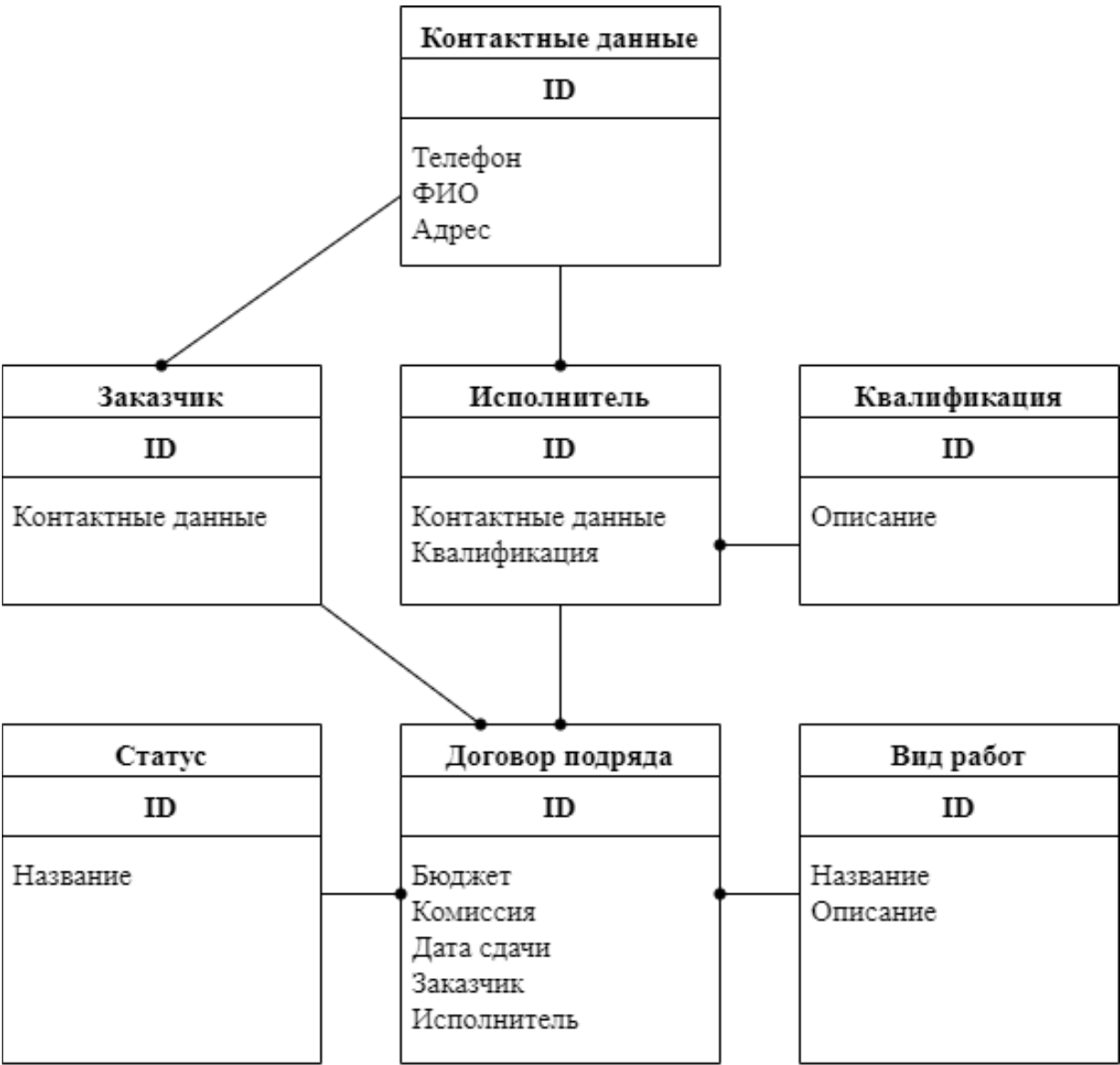


Рис. 2. Схема базы данных в нотации IDEF1X предметной области «Компания, занимающаяся перевозкой грузов»

Структура базы данных включают данные обо всех объектах в базе данных, таких как поля, таблицы и отношения. С помощью неё можно понять, как устроена база данных, как эффективно строить базу данных и управлять ею. В схеме структуры базы данных указываются типы данных, к которым принадлежат атрибуты каждой сущности.

Схема структуры базы данных:

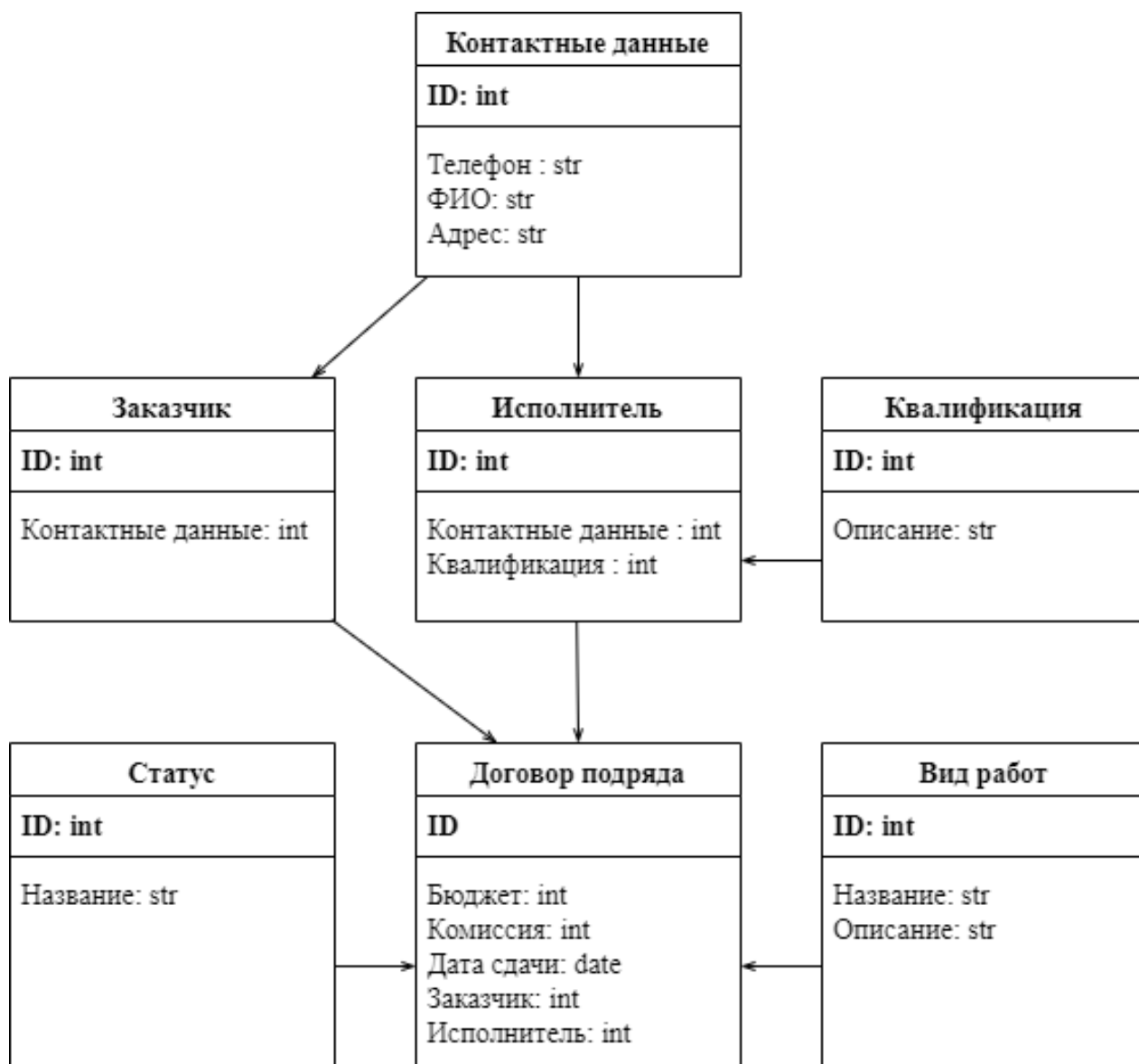


Рис. 3. Схема структуры базы данных предметной области «Компания, занимающаяся перевозкой грузов»

Описание столбцов таблиц базы данных:

Customer	ID	Целочисленный
	Contact Details	Целочисленный, ID ContactDetails
Contact Details	ID	Целочисленный
	FCS	Строка, ФИО
	Phone	Строка, номер телефона
	Address	Строка, адрес
Contract	ID	Целочисленный
	Customer	Целочисленный, ID Customer
	Executor	Целочисленный, ID Executor
	Budget	Целочисленный,

		количество д. е., выделенных на подряд
	Commission	Вещественный, процент из бюджета, который компания забирает себе
	EndDate	Дата, к которой подряд должен быть выполнен. Может быть NULL
Contract Status	ID	Целочисленный
	Name	Строка, текстовое описание статуса
Job	ID	Целочисленный
	Name	Строка, название вида работ
	Description	Строка, Описание, содержание работ
Executor	ID	Целочисленный
	ContactDetails	Целочисленный, ID ContactDetails
Qualification	ID	Целочисленный
	Description	Целочисленный, Описание квалификации

Табл. 1. Описание столбцов таблиц базы данных предметной области
«Компания, занимающаяся перевозкой грузов»

Таким образом, перед созданием базы данных необходимо располагать описанием выбранной предметной области, которое должно охватывать реальные объекты и процессы, иметь всю необходимую информацию для удовлетворения предполагаемых запросов пользователя и определить потребности в обработке данных. На основе такого описания на этапе проектирования базы данных осуществляется определение состава и структуры данных предметной области, которые должны находиться в базе данных и обеспечивать выполнение необходимых запросов и задач пользователя. Структура данных предметной области может отображаться информационно-логической моделью. На основе этой модели легко создается реляционная база данных.

В процессе выполнения лабораторной работы были получены навыки по созданию инфологической модели данных и структуры базы данных по заданной предметной области.