

Лекция 7

Проектирование БД. Этап концептуального проектирования.

Этапы проектирования

Проектирование БД

Концептуальное

Формулирование сущностей, атрибутов и связей

Логическое

Выбор модели данных и организация данных

Физическое



Концептуальное проектирование БД.

Или инфологическое,

Семантическое моделирование. Связано со смысловым содержанием данных, независимо от их представления в ЭВМ

На этом этапе создаются подробные модели пользовательских представлений данных предметной области. Затем они интегрируются в **инфологическую (концептуальную) модель**, которая фиксирует все элементы корпоративных данных подлежащих загрузке в БД

Инфологическая модель описывает структуру исследуемой предметной области. Она призвана выявить логико-семантические связи между данными. Эти связи – основа выбора модели данных и проектных решений относительно структуризации данных. Кроме того, логико-семантические связи нужны для определения ограничений целостности будущей базы данных.



Инфологическая модель

должна быть настолько глубокой, чтобы на ее основании можно было генерировать различные логические и физические модели данных.

С другой стороны, инфологическая модель должна включать такое формализованное описание предметной области, которое легко будет «читаться» не только специалистами по базам данных,

Ведь процесс проектирования длительный и трудоемкий и требует постоянного взаимодействия с заказчиком. Если БД является частью крупной корпоративной системы, то в этом процессе участвуют различные специалисты. Поэтому глубокое ясное и лаконичное изображение предметной области весьма полезно для уточнения деталей проекта и координации действий.

Чертеж выполненный по всем правилам понятен любому инженеру или рабочему

Описание алгоритма в виде блок-схемы может быть прочитано всяким программистом



Инфологическая модель

Проблема представления семантики давно интересовала разработчиков, и в семидесятых годах было предложено несколько моделей данных, названных семантическими моделями,

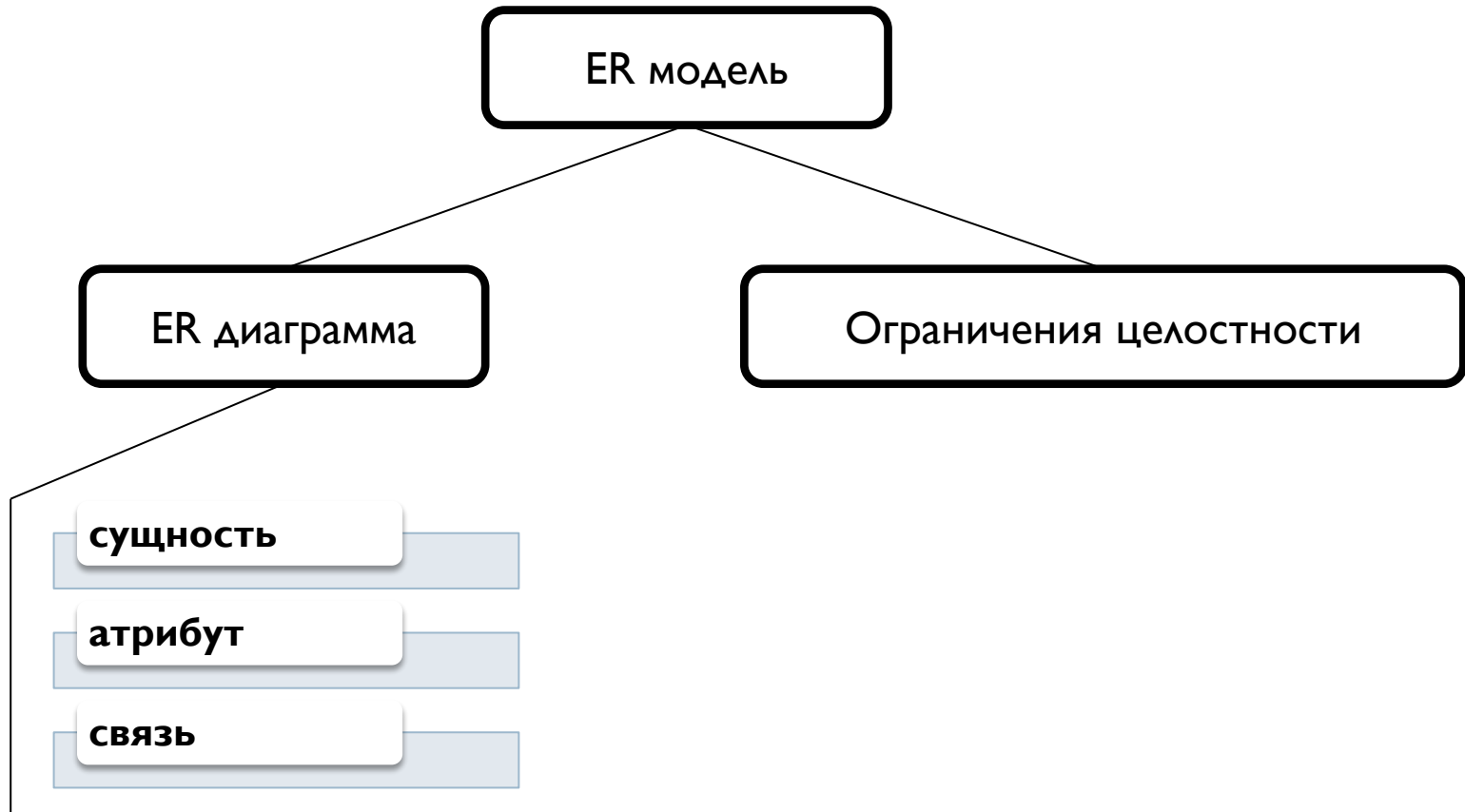
Среди них преобладают средства графического представления. Они чем-то напоминают язык блок-схем для описания алгоритмов или правила оформления конструкторской документации.

Наиболее простые и понятные обозначения имеет модель сущность-связь или “Entity Relationship model” (ER модель), предложенная *Ченом (Chen)* в 1976 году. Именно эта модель стала фактическим стандартом при инфологическом моделировании баз данных.



ER-модель

Инфологическая модель предметной области в терминах ER-модели имеет две составляющие



Сущность. Атрибуты.

с помощью **сущности** моделируется класс однотипных объектов предметной области

Сущность имеет **имя**, уникальное в пределах моделируемой системы. Так как сущность соответствует некоторому классу однотипных объектов, то предполагается, что в системе существует множество **экземпляров** данной сущности. Экземпляры сущности должны быть различны, т.е. у сущности должны быть свойства, уникальные для каждого экземпляра этой сущности

Объект, которому соответствует понятие сущности, имеет свой набор **атрибутов** — *характеристик, определяющих свойства данного представителя класса.*

При этом набор атрибутов должен быть таким, чтобы можно было различать конкретные экземпляры сущности.

Например, у сущности Сотрудник может быть следующий набор атрибутов: Табельный номер, Фамилия, Имя, Отчество, Дата рождения, Количество детей, Наличие родственников за границей.



Сущность. Атрибуты.

Набор атрибутов, однозначно идентифицирующий конкретный экземпляр сущности, называют *ключевым*.

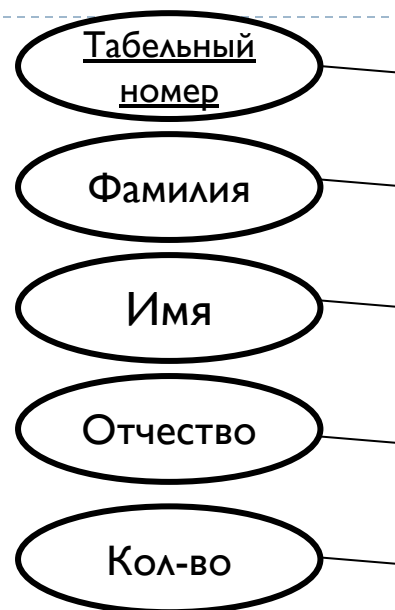
Для сущности *Сотрудник* ключевым будет атрибут *Табельный номер*, поскольку для всех сотрудников данного предприятия табельные номера будут различны. Экземпляром сущности *Сотрудник* будет описание конкретного сотрудника предприятия.



Сущность. Атрибуты.



A₁ A₂ A₃ A₄ A₅



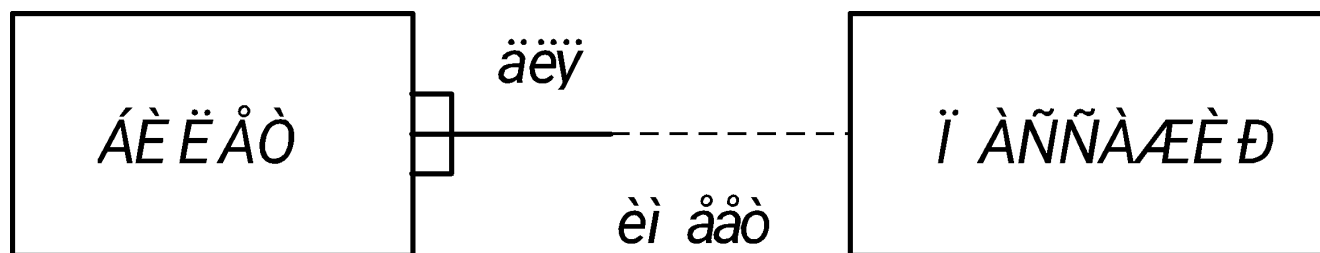
Идентификатор	Семантика	Домен
<u>A₁</u>	Табельный номер	N(6)
A ₂	Фамилия	C(30)
A ₃	Имя	C(30)
A ₄	Отчество	C(30)
A ₅	Кол-во детей	N(15)

СВЯЗЬ

Между сущностями могут быть установлены *связи* - *бинарные ассоциации*, показывающие, каким образом сущности соотносятся или взаимодействуют между собой. Связь может существовать между двумя разными сущностями или между сущностью и ей же самой (*рекурсивная связь*).

Она показывает, как связаны экземпляры сущностей между собой.

Связь имеет два конца. На каждом конце связи указывается его *имя*, *степень* (сколько экземпляров данной сущности связывается) и *обязательность* связи.

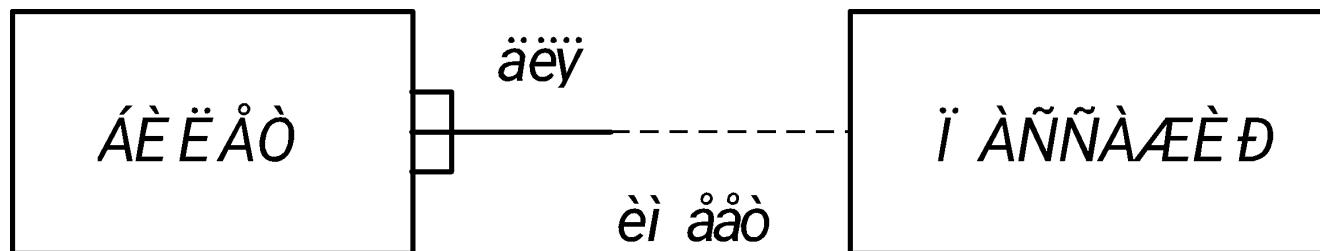


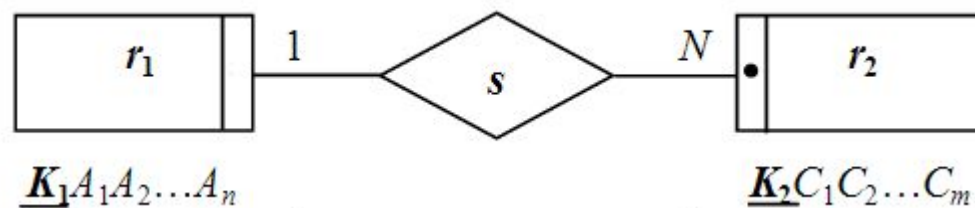
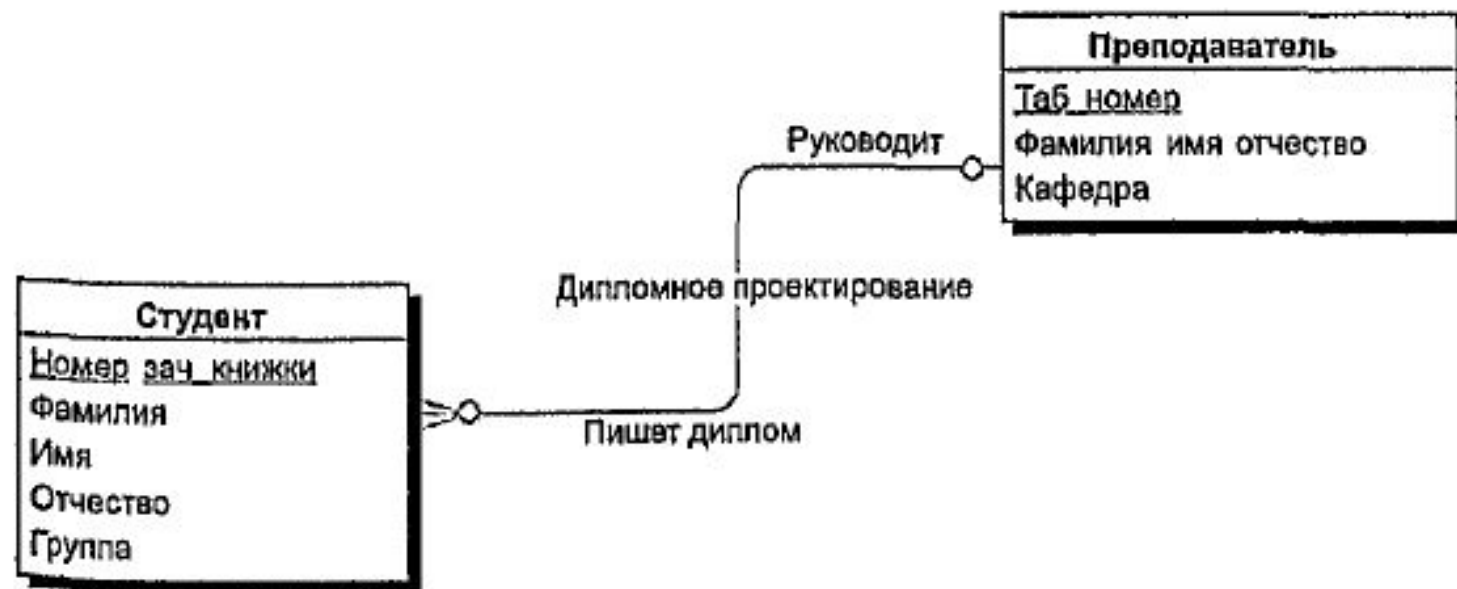
СВЯЗЬ

На приведенной схеме конец связи с именем *для* позволяет связывать с одним пассажиром несколько билетов, причем каждый билет должен быть обязательно связан с каким-либо пассажиром. Конец связи с именем *имеет* означает, что каждый билет принадлежит только одному пассажиру, причем пассажир может не иметь билета.

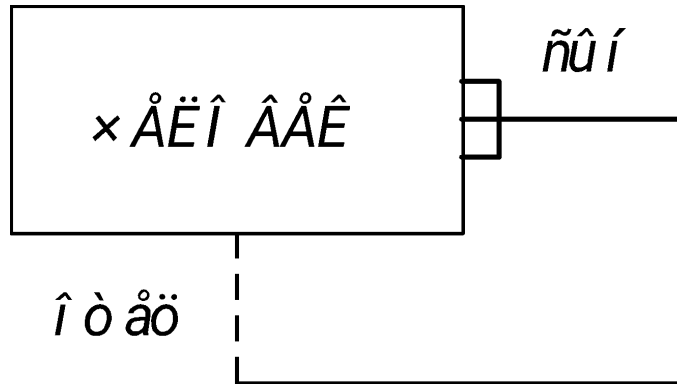
Лаконичная трактовка диаграммы:

- каждый билет предназначен для одного и только одного пассажира,
- каждый пассажир может иметь ноль или более билетов.





Рекурсивная связь. Пример.



В ER-схеме, включающей рекурсивную связь, конец связи с именем *сын* означает: у одного отца может быть более чем один сын. Смысл конца связи *отец*: не у каждого человека могут быть сыновья.

Краткая трактовка этой схемы:

- каждый человек является сыном одного и только одного человека,
- каждый человек может являться отцом для нуля и более людей.

Степень связи

Степень отражает однократность или многократность участия экземпляра в связи
Связи делятся на три типа по множественности:

- *один-к-одному* (1:1),
- *один-ко-многим* (1:M),
- *многие-ко-многим* (M:M).

Связь один-к-одному означает, что экземпляр одной сущности связан только с одним экземпляром другой сущности.

Связь 1: M означает, что один экземпляр сущности, расположенный слева по связи, может быть связан с несколькими экземплярами сущности, расположенными справа по связи.

связь «многие-ко-многим» (M:M) означает, что один экземпляр первой сущности может быть связан с несколькими экземплярами второй сущности, и наоборот, один экземпляр второй сущности может быть связан с несколькими экземплярами первой сущности.

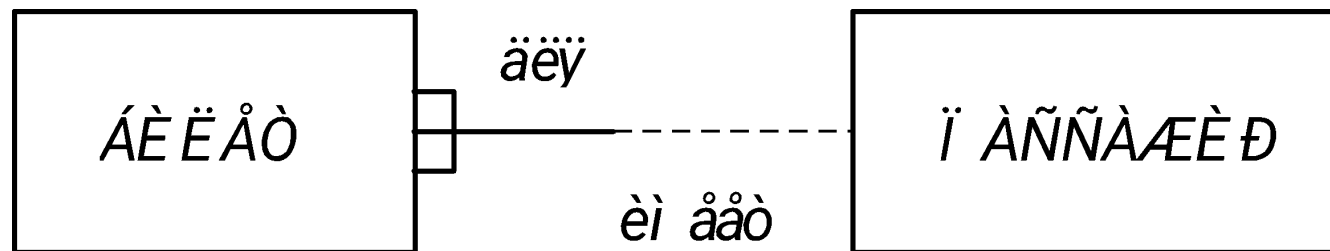
Класс принадлежности связи

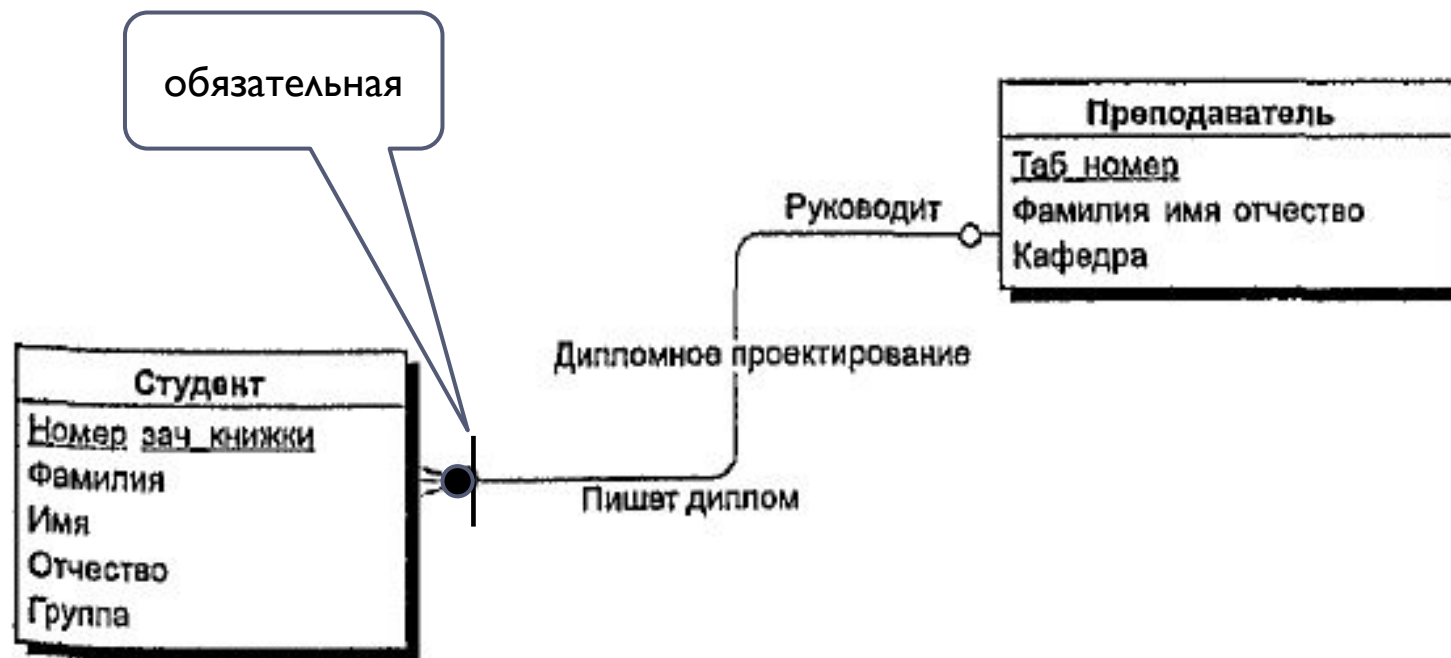
- **Обязательный.**

Соответствует модальности «должен». Это означает что каждый экземпляр сущности обязан быть связан с каким либо экземпляром другой сущности.

- **Необязательный**

- Соответствует модальности «может». Это означает что каждый экземпляр сущности может быть свободным от связи





Замечания

- Всякий атрибут может быть отнесен только к одной сущности или одной связи
- В результате построения модели предметной области в виде набора сущностей и связей получаем связный граф. В полученном графе необходимо избегать циклических связей — они выявляют некорректность модели.
- Определения конструктивных элементов не отличаются особой конкретностью. Поэтому два разработчика, моделируя одну и ту же предметную область, могут получить различные *ER*-диаграммы

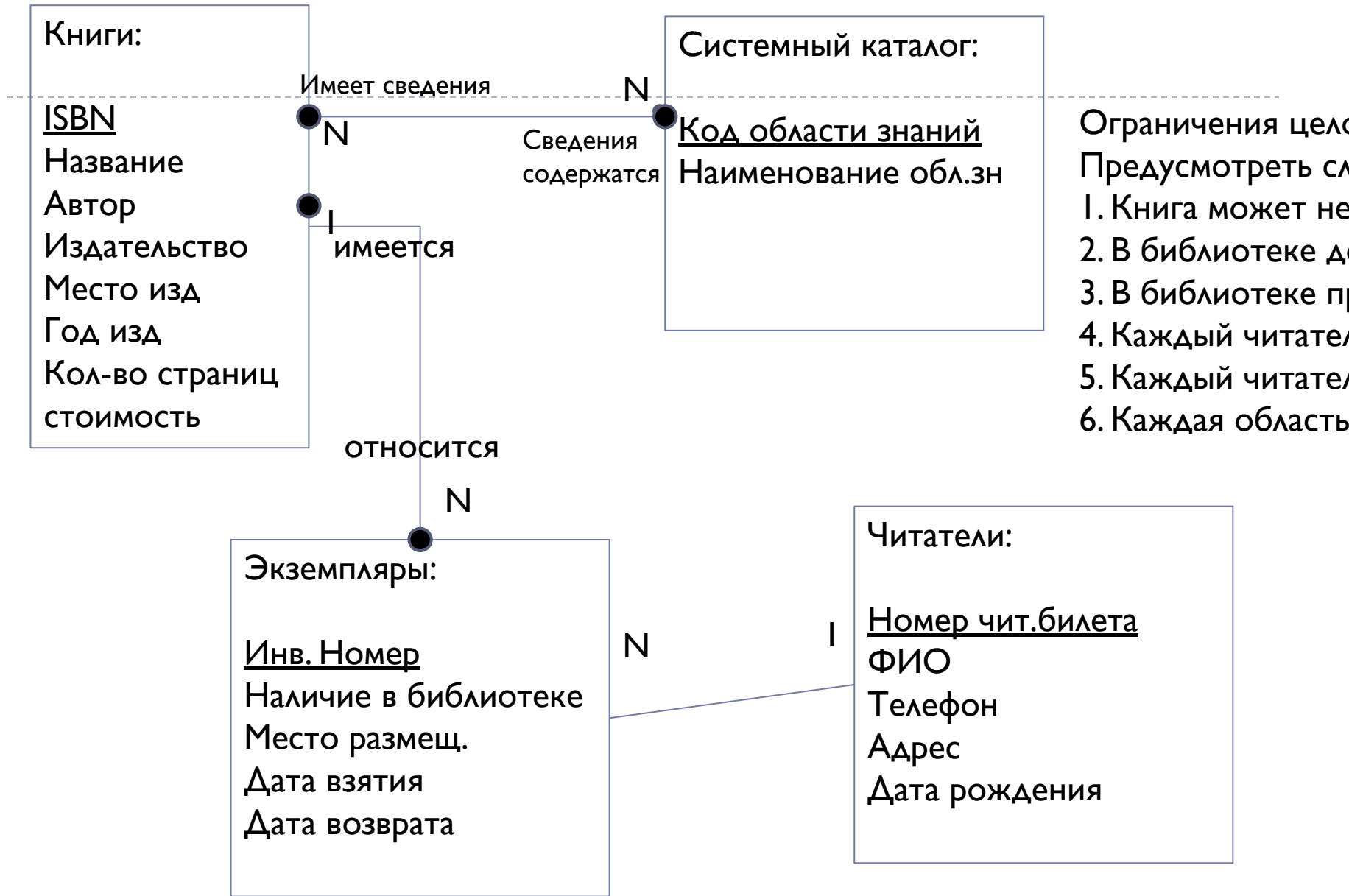


Пример

- В качестве примера спроектируем инфологическую модель системы, предназначенной для хранения информации о книгах и областях знаний, представленных в библиотеке.
- Описание предметной области приведено в отдельном файле.
- Разработку модели начнем с выделения основных сущностей.



ER-модель



Пример. Описать ER диаграмму следующей предметной области. (самост.)

Объектом моделирования (предметной областью) выступает работа предприятия по ремонту.

Назначение базы данных:

- учет занятости мастеров
- расчет оплаты труда мастеров

В процессе анализа предметной области был установлен набор атрибутов, значения которых следует хранить в базе данных, и выявлены связи между атрибутами

Идентификатор	Семантика	Домен
Tn	Табельный номер мастера	N(6)
F	Ф.И.О.	C(30)
R	Номер квалификационного разряда мастера	N(2)
U	Название улицы	C(20)
D	Номер дома	C(4)
P	Площадь ремонта, м ²	N(7.1)
C	Категория сложности ремонта	N(3.1)
Tr	Тариф (оплата за 1 м ² ремонта согласно разряду)	N(5.1)
Data	Дата работы мастера	D
V	Объем работы, выполненной мастером за день, м ²	N(5.1)

Особенности предметной области и специфика связей между атрибутами:

- каждый дом идентифицируется названием улицы и номером дома,
- табельный номер однозначно определяет значения всех атрибутов мастера,
- учитывать следует только те дома, которые подлежат ремонту,
- номер разряда задает значение тарифа,
- каждый рабочий день мастер может работать на ремонте только одного дома,
- один и тот же дом могут одновременно ремонтировать несколько мастеров,
- некоторые мастера могут быть временно без работы,
- зарплата мастера за месяц начисляется (по сдельной форме) согласно формуле

$$Tr \cdot (\sum V_i \cdot C_i),$$

где суммирование берется по всем дням работы мастера в заданном месяце, а значение Tr соответствует разряду мастера





