

Моделирование данных

А3. Физические модели данных



Московский государственный технический университет
имени Н.Э. Баумана

Факультет ИБМ

сен 2024 года

Москва

Артемьев Валерий Иванович © 2024

Курс «Моделирование данных»

3. Физические модели данных

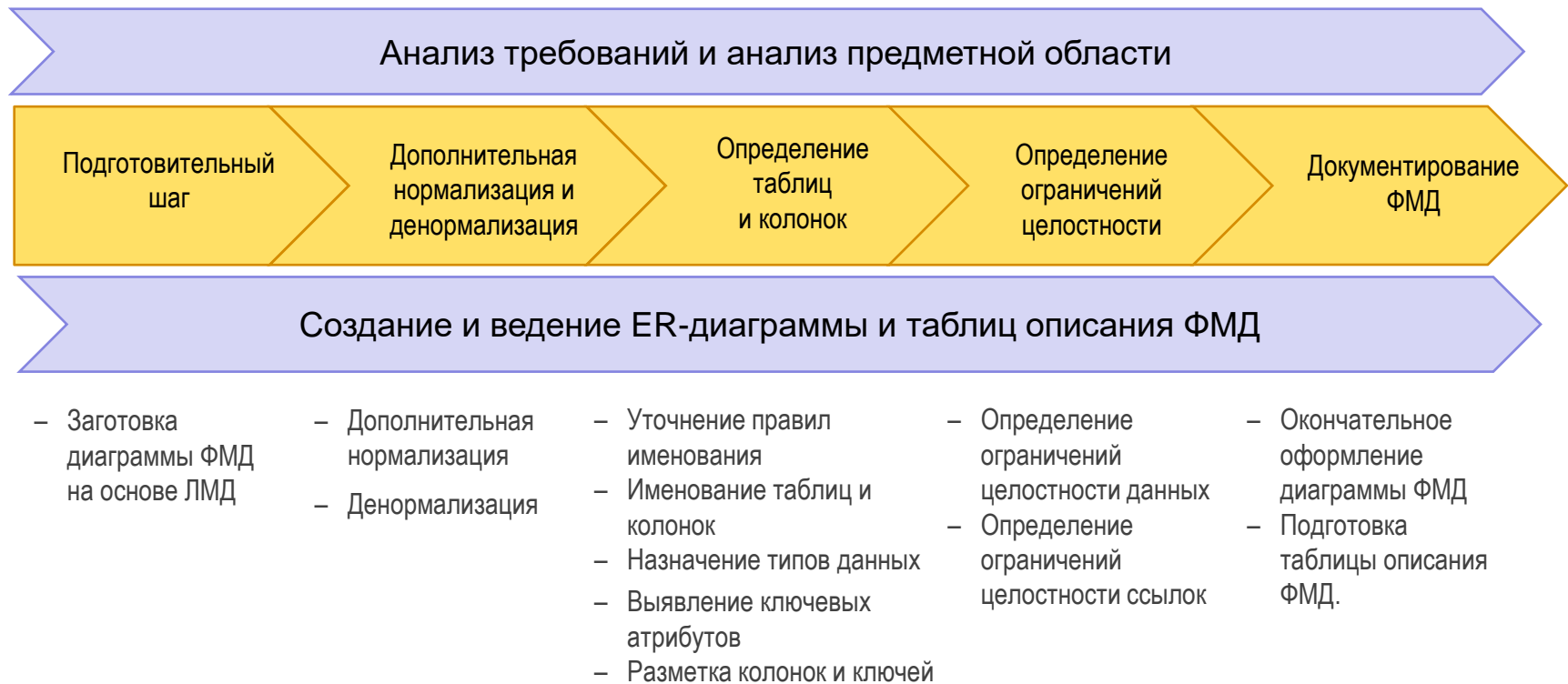
- Определение физической модели данных
 - Шаги создания физической модели данных
 - Подготовительный шаг
-
- Именованние таблиц и колонок ФМД
 - Типы данных СУБД
 - Необязательные значения NULL
 - Выбор имён для таблиц, колонок примера
 - Определение типов данных БД примера
 - Целостность данных и бизнес-правила
 - Определение процедур контроля целостности данных для примера
-
- Первичные и внешние ключи (РК и FK)
 - Объявление ключевых атрибутов примера
 - Целостность ссылок
 - Определение процедур контроля целостности ссылок для примера
-
- Нотации физической модели данных
 - Создание диаграммы «сущность-связь» для ФМД примера
 - Оформление таблицы описания физической модели данных примера
 - Описание данных – в комментарии к таблицам и атрибутам

Физическая модель данных

Детальная модель данных отражает необходимые потребности бизнеса, ограничения выбранной абстракции и конкретной реализации в виде БД.

- Выполняется *дополнительная нормализация или денормализация баз данных.*
- Сущности данных представляются в виде *таблиц и представлений базы данных.*
- Атрибуты реализуются в виде *колонок таблицы БД.*
- Идентификаторы и ссылки преобразуются в *первичные (primary key) и внешние ключи (foreign key) БД.*
- Использование *типов данных SQL.*
- На основе физической модели БД формируется *DDL-скрипты для генерации объектов БД .*

Шаги создания физической модели данных



Подготовительный шаг

Необходимо скопировать ER-диаграмму ЛДМ:

- Открыть диаграмму ЛДМ и сохранить её как диаграмму ФМД для дальнейшей работы.

Дальнейшие действия зависят от нотации ER-диаграммы ЛДМ в графических редакторах:

- Наиболее просто создать ФМД из ЛМД для упрощенной нотации в MS PowerPoint.
- Создание ФМД в других нотациях может потребовать переименования атрибутов или пересоздания блоков сущностей в худшем случае.

Именованние таблиц и колонок ФМД

- Именуйте таблицы и колонки ФМД на основе имён сущностей и атрибутов ЛМД
 - Сохранять выбранный порядок слов в именах
 - Русские сокращения применять, но не переводить
 - Формат имени змейка (snake_case): **creator_role**
-
- Есть повод поговорить о правилах именования
 - Правила именования <https://www.sqlstyle.guide/ru/>
 - Книга «Стиль программирования Джо Селко на SQL»
 - Стандарт именования элементов данных ГОСТ Р ИСО/МЭК 11179-1-2010

Правила именования элементов данных имеют более широкую область применения в ИТ.



Стандарт именования данных ГОСТ Р ИСО/МЭК 11179-1-2010

Семантика имён

- Имена таблиц и колонок имеют следующие составляющие:
 - Термы классов объектов (сущностей)
 - Термы свойств объектов
 - Термы представления
 - Уточняющие термы
- Должны быть ровно один терм классов, один терм свойств и может быть один терм представления

Пример: **Общая Сумма Стоимости** за **Бюджетный Период**

Синтаксис имён

- Терм класса должен занимать 1-ое место в имени
- Уточняющие термы должны предшествовать уточняемым термам
- Терм свойства должен занимать следующую позицию
- Терм представления занимает последнее место

Пример: **Стоимость_БюджетныйПериод_Общая_Сумма**
Cost_BudgetPeriod_Total_Amount

Преобразования имён ЛМД в имена ФМД

- Варианты правил именования
 - *по-английски*, нужно переводить и вести словарь
 - *транслитерация* с настройкой <http://translit-online.ru/>
 - *русский язык* поддерживается в некоторых БД
- Преобразования имён атрибутов (сущностей) в имена колонок (таблиц)

Кто знает,
что такое
транслитерация?

Формат имени	Имя атрибута (сущности)	Транслитерация		Перевод	Имя в кавычках
		Прямой порядок слов	Обратный порядок слов	Обратный порядок слов	
Прямой порядок слов	имя_издателя	imja_izdatelja			
Обратный порядок слов	издатель_имя		izdatel_imja	publisher_name	
Имя в кавычках	“имя издателя”				“имя издателя”

- Стандарт транслитерации ГОСТ 7.79-2000

Основные классы типов данных

Тип данных – именованная категория множества значений и характеристик элемента данных, а также набор допустимых операций над ними.



*Какие типы данных вы знаете
или сталкивались на практике
в программировании?*

Основные классы типов данных

Тип данных – именованная категория множества значений и характеристик элемента данных, а также набор допустимых операций над ними.



- Символьные типы данных
- Числовые типы данных
- Типы данных даты и времени
- Типы больших объектов данных
- Логические типы данных
- Экзотические типы данных

ФМД содержит *типы данных в реляционных БД*
ЛМД могут содержать *математические или
прикладные типы данных*.

Числовые и логические типы данных

Тип данных	Oracle	SQL Server	MySQL	PostgreSQL	
SMALLINT	Y	Y	Y	Y	
INTEGER	Y	Y	Y	Y	
INT	Y	Y	Y		
BIGINT		Y	Y	Y	
NUMERIC	Y	Y	Y	Y	
DECIMAL					
MONEY		Y		Y	
REAL	Y	Y	Y	Y	
FLOAT	Y	Y	Y		
DOUBLE PRECISION	Y	Y	Y	Y	
BIT		Y	Y		

Символьные типы данных

Тип данных	Oracle	SQL Server	MySQL	PostgreSQL	
SMALLINT	Y	Y	Y	Y	
INTEGER	Y	Y	Y	Y	
INT	Y	Y	Y		
BIGINT		Y	Y	Y	
NUMERIC	Y	Y	Y	Y	
DECIMAL					
MONEY		Y		Y	
REAL	Y	Y	Y	Y	
FLOAT	Y	Y	Y		
DOUBLE PRECISION	Y	Y	Y	Y	
BIT		Y	Y		

Типы данных даты и времени

Тип данных	Oracle	SQL Server	MySQL	PostgreSQL	
SMALLINT	Y	Y	Y	Y	
INTEGER	Y	Y	Y	Y	
INT	Y	Y	Y		
BIGINT		Y	Y	Y	
NUMERIC	Y	Y	Y	Y	
DECIMAL					
MONEY		Y		Y	
REAL	Y	Y	Y	Y	
FLOAT	Y	Y	Y		
DOUBLE PRECISION	Y	Y	Y	Y	
BIT		Y	Y		

Типы больших объектов данных

Тип данных	Oracle	SQL Server	MySQL	PostgreSQL	
SMALLINT	Y	Y	Y	Y	
INTEGER	Y	Y	Y	Y	
INT	Y	Y	Y		
BIGINT		Y	Y	Y	
NUMERIC	Y	Y	Y	Y	
DECIMAL					
MONEY		Y		Y	
REAL	Y	Y	Y	Y	
FLOAT	Y	Y	Y		
DOUBLE PRECISION	Y	Y	Y	Y	
BIT		Y	Y		

Объявление необязательных значений

NULL

NOT NULL

Примеры

Выбор имён таблиц и колонок примера

Объявление ключей

Первичный ключ PK
PRIMARY KEY

Альтернативные ключи
UN
UNIQUE

Внешние ключи
FK
FOREIGN KEY

Процедуры проверки уникальности и целостности ссылок

Описание данных – комментарии

Комментарии к таблицам

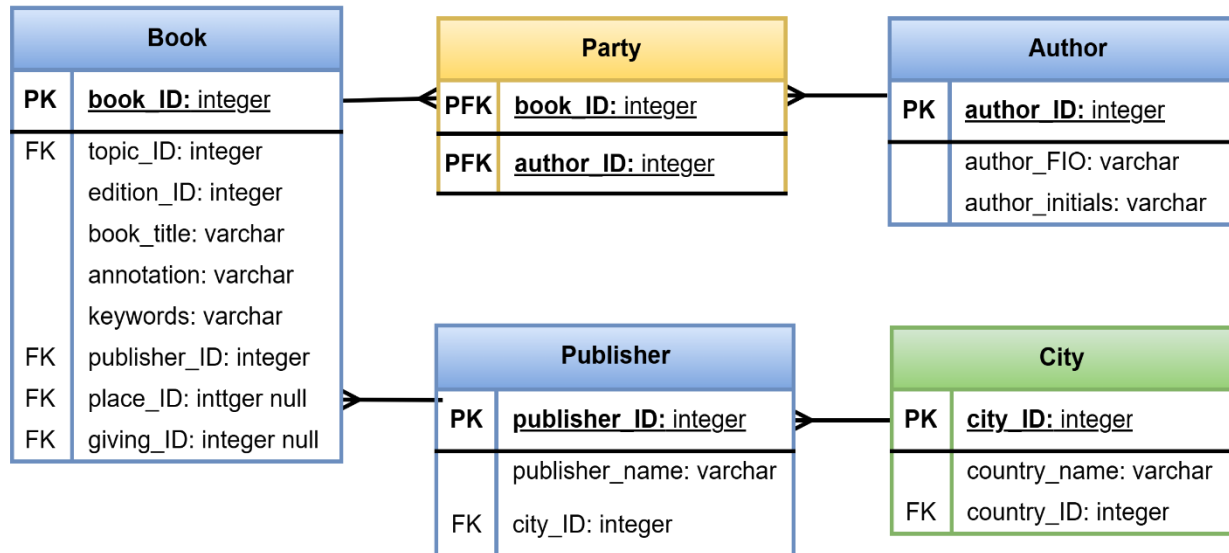
Комментарии к колонкам

Описание данных – комментарии

Комментарии к таблицам

Комментарии к колонкам

Нотация ER-диаграммы физической модели данных



Нотация Р. Баркера

- Имя таблицы/ представления существительное латиницей в единственном числе на полочке или подчёркнуто.
- В блоке задаются имена колонок латиницей без пробелов, без повторения или с повтором имени сущности.
- Следует указывать (PK) для идентификатора и (FK) для ссылки.
- Указывать (null) для необязательных атрибутов.
- Допустимо указывать прикладные типы данных для колонки через «:».

Состав таблицы описания физической модели данных

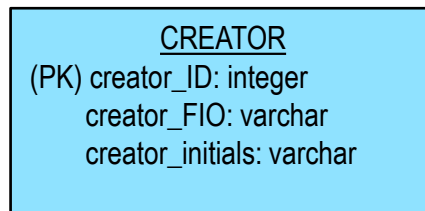
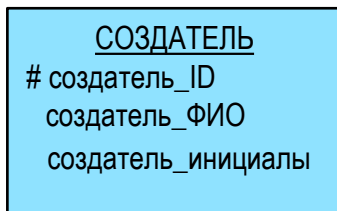
- Имя базы данных (имя предметной области)
- Вид набора данных: таблица/ представление
- Имя таблицы (сущности)
- Имя колонки (атрибута)
- Описание колонки (атрибута)
- Обязательность значения: NULL или NOT NULL
- Имя домена данных и его описание
- Тип данных и *длина/ разрядность*
- Процедуры целостности данных и ссылок
(допустимые ограничения типа, правила целостности, первичные и внешние ключи)
- [Примеры значения данных]

Пример: Домашняя библиотека

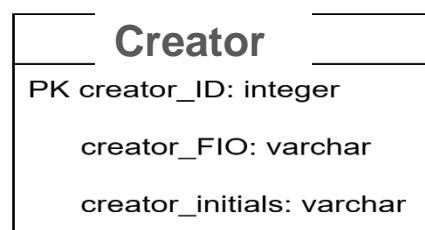
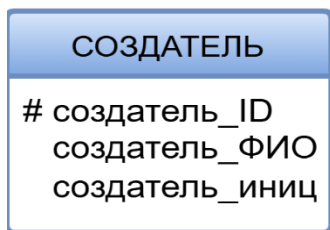
Создать каталог домашней библиотеки

- **Атрибуты поиска книги:**
 - Тематика, вид издания
 - Название издания и/ или ключевые слова
 - Фамилию, имя (инициалы) автора, редактора, составителя, переводчика или художника
 - Название и место издательства
 - Год выпуска.
- **Результаты поиска:**
 - Хранимые атрибуты книги (включая аннотацию, описание приложения и сведения об оригинале перевода)
 - Место хранения издания (шкаф, полка)
 - Текущий держатель (читатель) издания, его контакты
- **В библиотеке каждая книга присутствует в одном экземпляре**
- **Расширения видов изданий:** электронные книги и аудиокниги, хранимые на компьютерах, ноутбуках, планшетах и смартфонах
- **Расширение мест хранения:** дача, офис

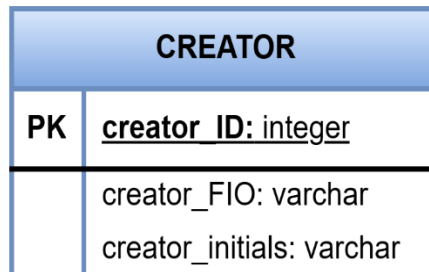
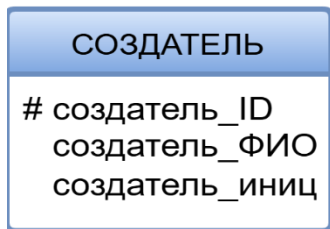
Преобразование блоков ER-диаграммы



PowerPoint
Замена и редактирование
названий атрибутов



Draw.io
Замена и редактирование
названий атрибутов



Draw.io
Замена блоков,
ввод и копирование
названий атрибутов

Типы и домены данных

■ Что такое домен данных?

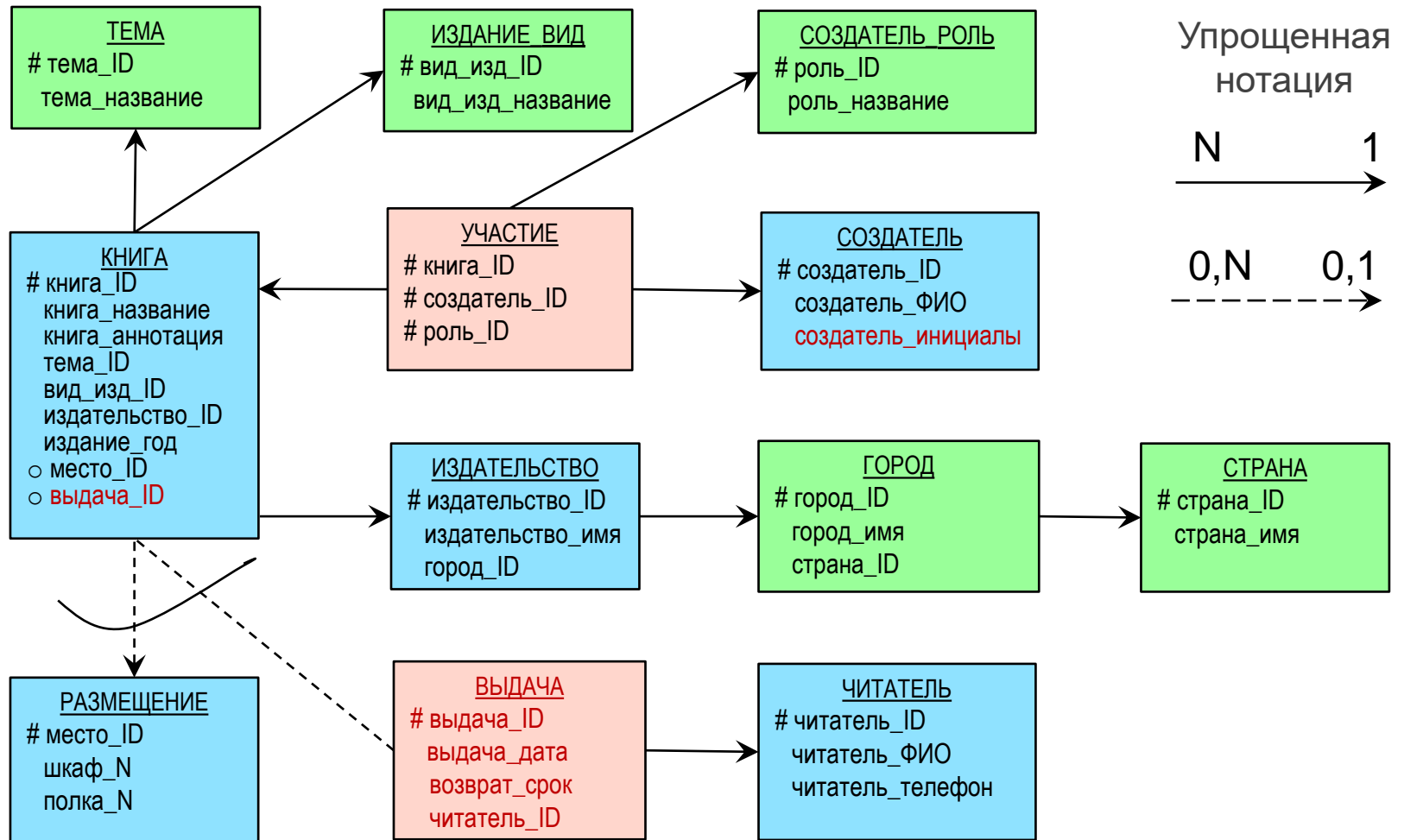
- Именованная область допустимых значений однотипных атрибутов – прикладной тип данных, задаёт ограничения для значений
- Тип данных
- Диапазон значений
- Список допустимых значений
- Условие

Домен данных – тип данных и его ограничения, задает область допустимых значений в виде длины/разрядности, диапазона или списка значений, формата (шаблона) значения.

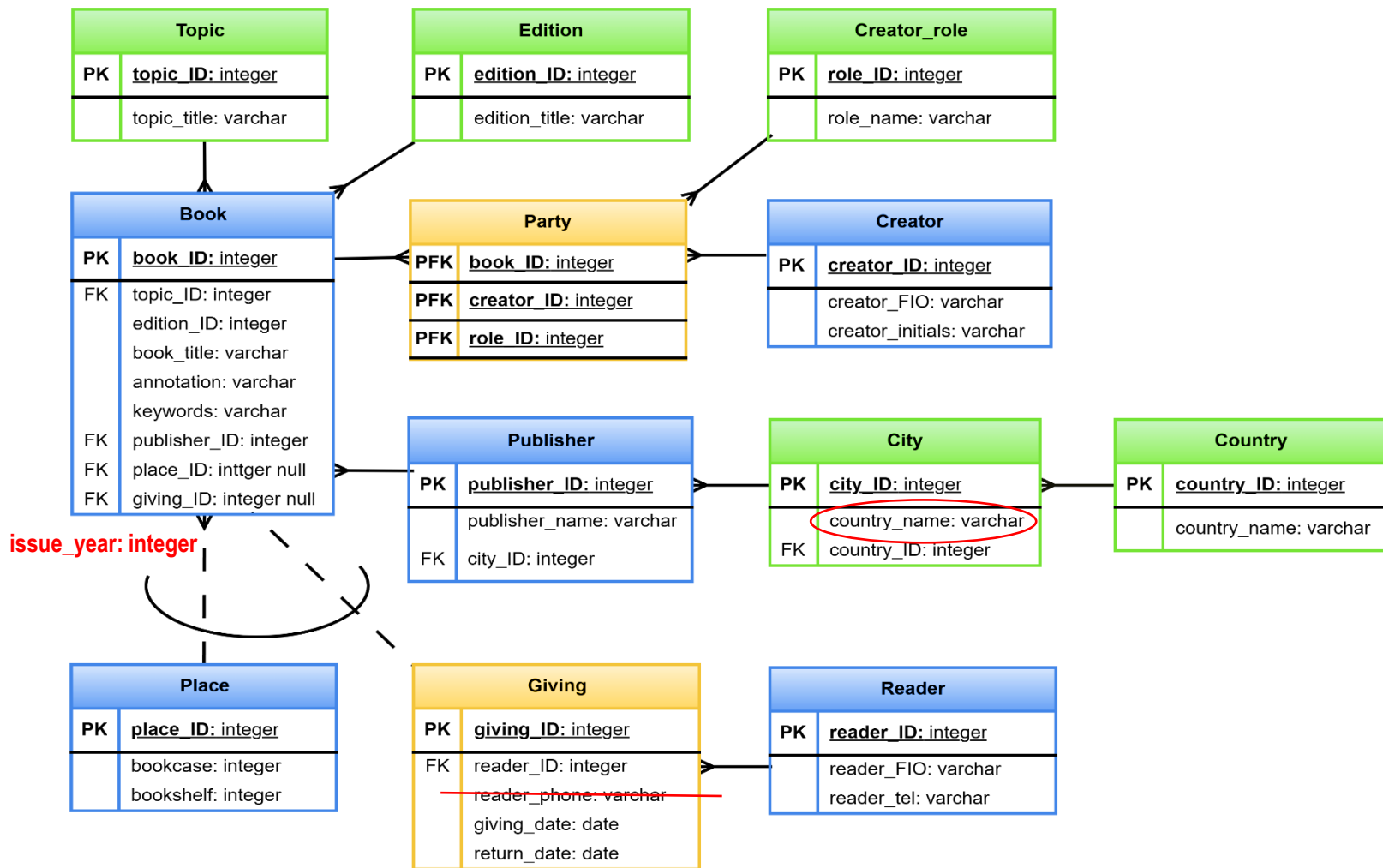
■ Примеры доменов

- Тип_ИНН_юрлиц или Цифры(10)
- Деньги(20,4)
- Список: “малый”; “средний”; “большой”

Логическая модель данных «Домашняя библиотека»



Физическая модель данных «Домашняя библиотека» (draw.io)



Описание физической модели данных «Домашняя библиотека»

**Терпения и удачи всем, кто связан
с моделированием данных**

Спасибо за внимание!

Валерий Иванович Артемьев

**Департамент данных, проектов и процессов
Банк России**

Тел.: +7(495) 753-96-25

e-mail: avi@cbr.ru