

— — — — — ЛЕКЦИЯ 2 — — — — —

СЛОВАРЬ ДАННЫХ

РАЗБОР ЛАБОРАТОРНОЙ РАБОТЫ 1

Вопрос 1



Вопрос 2



Слайд 3 из 21

Словарь данных

Словарь данных представляет собой централизованное хранилище подробной информации об используемых в приложении сущностях данных

Информация словаря данных включает в себя определение состава *структур данных* предметной области, заданных или разрешенных значений, типов данных, длины и формата *элементов данных*, из которых состоят эти структуры

Базовые понятия

Структура — совокупность элементов данных

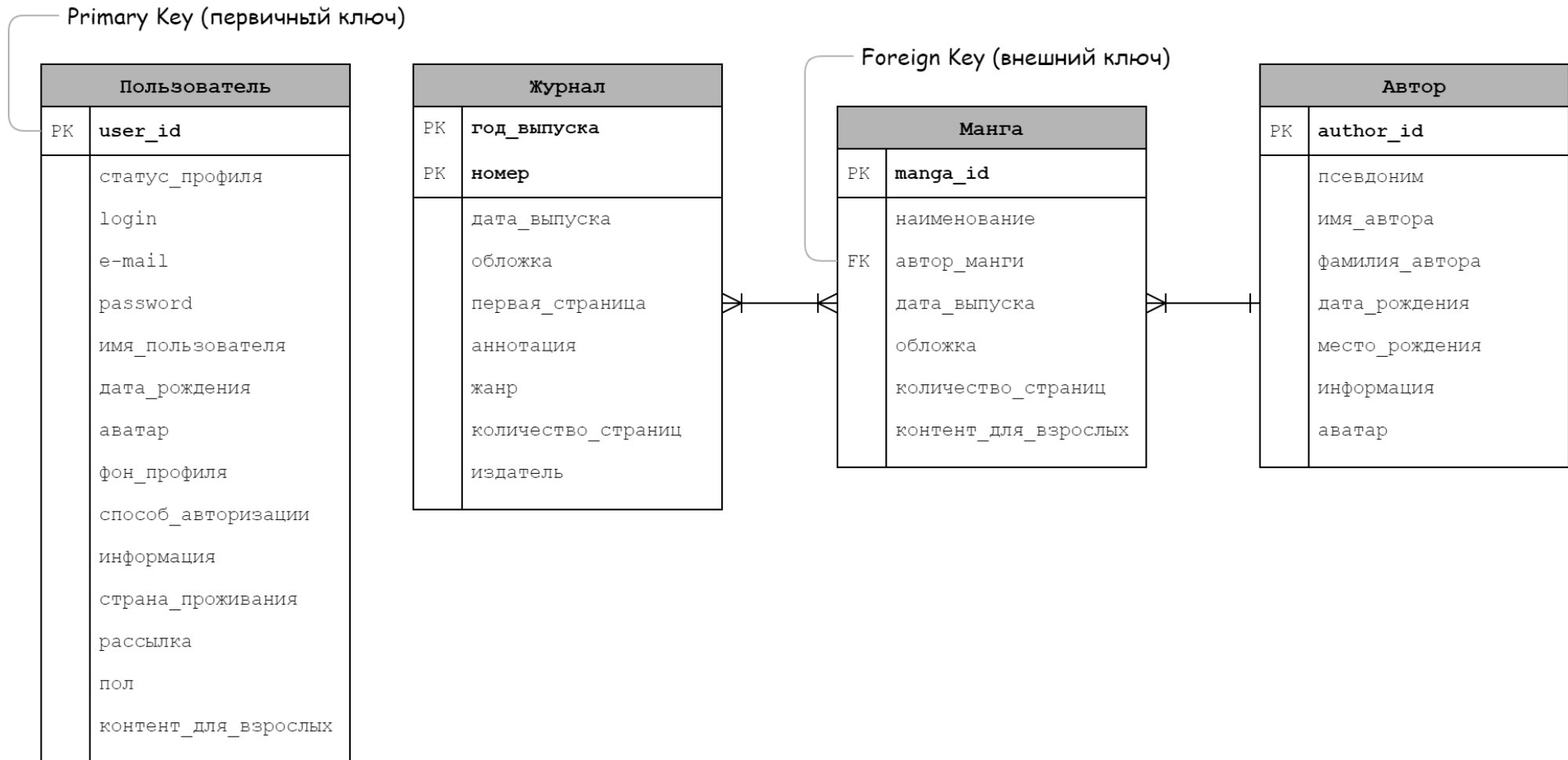
Простейший элемент данных — дальнейшая декомпозиция которого невозможна или не нужна

В описании простейших элементов указывают тип, длину, диапазон числовых значений, список разрешенных значений и другие уместные атрибуты

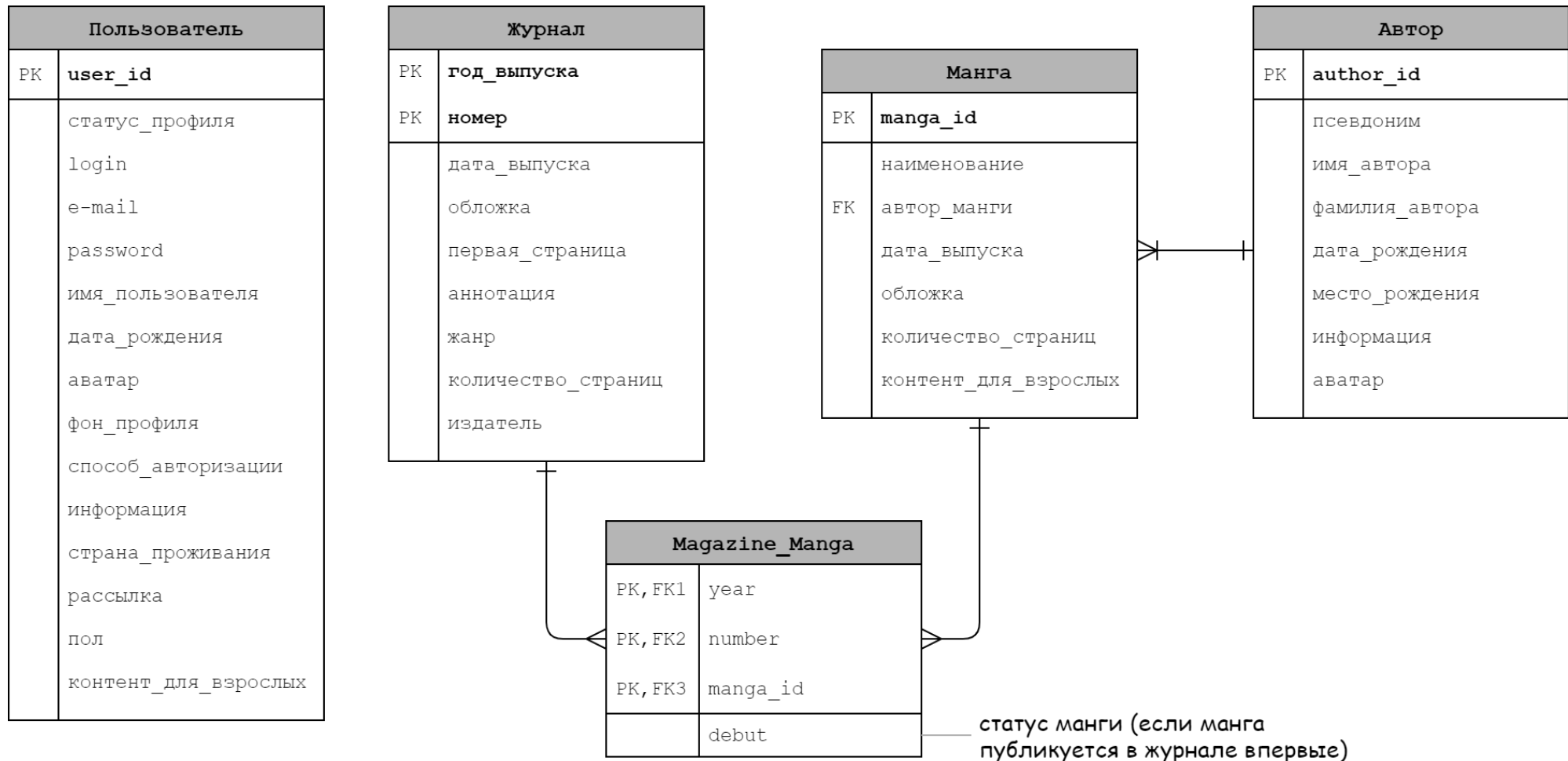
Необязательные элементы, значение которых не должно предоставляться пользователем или системой, заключают в скобки

Выполнение ЛР 1

Шаг 1 — Фрагмент даталогической модели



Шаг 2 — Реструктуризация связи М : М



Шаг 3 — Словарь данных

Структура или элемент данных	Тип данных	Длина	Значение
Пользователь			
Идентификатор	INT	10	Первичный ключ — автоинкрементный номер записи, генерируемый системой, начиная с 1
Статус профиля	VARCHAR	5	ADMIN USER RCF (registration confirmation) — статус для учетной записи, которая требует подтверждения по email
Логин	VARCHAR	25	Может содержать только символы латинского алфавита, подчеркивание и цифры
E-mail	VARCHAR	50	Должен соответствовать стандарту RFC 5322
Пароль	VARCHAR	50	Может содержать символы латинского алфавита, числа и символы из следующего после двоеточия списка: ! @ # \$ % ^ & ? * _ Остальные символы, включая пробел, запрещены

Шаг 3.1 — Таблица «Пользователь»

Имя пользователя	VARCHAR	50	Может содержать все буквенно-цифровые символы, включая символы национального алфавита
Дата рождения	DATE	—	Используется для рекомендаций
(Аватар)	VARCHAR	255	Содержит путь файла
(Фон профиля)	VARCHAR	255	Содержит путь файла
Способ авторизации	VARCHAR	20	Login Google Twitter Facebook
(Информация)	TEXT	—	Содержит дополнительную информацию, которую пользователь указывает о себе
(Страна проживания)	ENUM	—	В интерфейсе выбирается из выпадающего списка
Рассылка	BIT	1	Поле не может быть пустым 0 — Нет, спасибо 1 — Да (по умолчанию)

Шаг 3.1 — Таблица «Пользователь» (продолжение)

(Пол)	BIT	1	Используется для рекомендаций / рекламы манги 0 — Женщина 1 — Мужчина NULL (то есть отсутствие значения) соответствует: Не указан или не определен (по умолчанию)
(Контент для взрослых)	BIT	1	0 — Нет, я не хочу видеть «Контент для взрослых» 1 — Да, всегда отображать «Контент для взрослых» NULL (отсутствие значения) соответствует: Всегда спрашивать разрешение на отображение контента для взрослых (по умолчанию)

Шаг 3.2 — Таблица «Номер журнала»

Номер журнала			
Год выпуска	YEAR	4	Атрибут составного первичного ключа — содержит значение года, за которым закреплен выпуск данного номера журнала
Номер	TINYINT	2	Атрибут составного первичного ключа — содержит порядковый номер журнала за определенный год (не может превышать двухзначное число)
Дата выпуска	DATE	—	Содержит дату фактического выпуска журнала, которая может отличаться от года, за которым закреплен выпуск номера журнала
Обложка	VARCHAR	255	Содержит путь файла
Первая страница	VARCHAR	255	Содержит путь файла
Аннотация	VARCHAR	255	Содержит короткое описание содержания номера журнала

Шаг 3.2 — Таблица «Номер журнала» (продолжение)

Жанр	ENUM	—	Возможные значения: Shonen; Shojo; Seinen; Josei (используется для рекомендаций, в интерфейсе может не отображаться)
Количество страниц	SMALLINT	4	Не может превышать 620 страниц (ограничение типографского издания)
Издатель	VARCHAR	50	Значение по умолчанию: Kodansha (по умолчанию указано основное издательство)

Шаг 3.3 — Таблица «Манга»

Манга			
Идентификатор	INT	10	Первичный ключ — автоинкрементный номер записи, генерируемый системой, начиная с 1
Наименование	VARCHAR	255	
Автор манги	INT	10	Содержит идентификатор автора
Дата выпуска	DATE	—	
Обложка	VARCHAR	255	Содержит путь файла
История	VARCHAR	255	Содержит небольшое описание по типу логлайна (то есть краткая аннотация к манге, которая передает суть истории и ее основную драматическую коллизию)
Количество страниц	SMALLINT	5	Если манга состоит из глав, то здесь указывается суммарное количество страниц
Контент для взрослых	BIT	1	Поле не может быть пустым 0 — Отсутствует 1 — Могут присутствовать сексуальные сцены, эпизоды с употреблением наркотиков, нецензурной бранью, насилием, et cetera.

Шаг 3.4 — Связующая таблица «Magazine_Manga»

Magazine_Manga			
Год выпуска	YEAR	4	Составной внешний ключ — ссылка на часть составного первичного ключа «Год выпуска» таблицы «Номер журнала»
Номер	TINYINT	2	Составной внешний ключ — ссылка на часть составного первичного ключа «Номер» таблицы «Номер журнала»
Идентификатор манги	INT	10	Внешний ключ — ссылка на первичный ключ таблицы «Манга»
Дебют	BIT	1	Поле не может быть пустым 1 — Да, манга публикуется впервые 0 — Нет, манга уже публиковалась

Шаг 3.5 — Таблица «Автор»

Автор			
Идентификатор	INT	10	Первичный ключ — автоинкрементный номер записи, генерируемый системой, начиная с 1
(Псевдоним)	VARCHAR	50	Может содержать все буквенно-цифровые символы
Имя автора	VARCHAR	50	Может содержать все буквенно-цифровые символы Вместо имени автора может быть указано наименование художественного коллектива
(Фамилия автора)	VARCHAR	50	Может содержать все буквенно-цифровые символы
(Дата рождения)	DATE	—	В системе можно задать отображение только числа и месяца рождения
(Место рождения)	VARCHAR	255	

Шаг 3.5 — Таблица «Автор» (продолжение)

(Информация)	TEXT	—	Краткая информация об авторе (по согласованию с автором) Подсказка при заполнении: «По умолчанию должна содержать строку следующего вида (пример): Дебютировал в 1988 году отдельным томом «It is awfully hard work doing nothing» (Kodansha)»
(Аватар)	VARCHAR	255	Содержит путь файла

ПРИЛОЖЕНИЕ А — Принятые типы данных в SQL (неполный список)

1	TINYINT	Целочисленный тип размером 1 байт Со знаком от -128 до 127, без знака от 0 до 255
2	SMALLINT	Целочисленный тип размером 2 байта Со знаком от -32 768 до 32 767, без знака от 0 до 65 535
3	MEDIUMINT	Целочисленный тип размером 3 байта Со знаком от -8 388 608 до 8 388 607, без знака от 0 до 16 777 215
4	INT	Целочисленный тип размером 4 байта Со знаком от -2 147 483 648 до 2 147 483 647, без знака от 0 до 4 294 967 295
5	BIGINT	Целочисленный тип размером 8 байт Со знаком от -2^{63} до $2^{63}-1$, без знака от 0 до $2^{64}-1$
6	DECIMAL	Тип с фиксированной точкой (часто используется для денежных данных) DECIMAL (size, d), где size — общее количество цифр (максимум 65), d — количество цифр после точки (максимальное значение для d — 30). Значения по умолчанию — 10 (для size) и 0 (для d). Например, DECIMAL (7, 2) может хранить значения от -99999.99 до 99999.99

7	FLOAT	<p>Тип с плавающей точкой размером 4 байта</p> <p>Обратите внимание, что каждая база данных имеет разные реализации типа FLOAT. В старых версиях MySQL (до 8.0.17) можно было указать точность для числа с плавающей запятой (аналогично DECIMAL), то есть FLOAT (size, d). В текущих версиях данный тип выражается как FLOAT (n), где n определяет, будет ли значение сохранено как FLOAT или преобразовано в DOUBLE. При n от 0 до 23 значение хранится в виде 4-байтового столбца с одинарной точностью, при n от 24 до 53 в виде 8-байтового столбца с двойной точностью (тип DOUBLE). По умолчанию значение n равно 53 (двойная точность).</p> <p>Значения с плавающей точкой представляются в форме экспоненциальной записи: $N = M \cdot n^p$, где N — число, M — мантисса, n — основание, p — порядок. Например, число $N = 0.004137 = 4.137 \cdot 10^{-3}$ (4,137 — мантисса, 10 — основание, -3 — порядок). Эта запись для человека, а для хранения в памяти машины удобна запись в форме MEp, где M — мантисса, E — «$\cdot 10^{\wedge}$» (умножение на 10 в степени), p — порядок. Таким образом, $4.137 \cdot 10^{-3}$ это 4.137E-3, а $4.137 \cdot 10^3$ это 4.137E+3.</p> <p>Диапазон значений для одинарной точности: от -3.40E+38 до -1.18E-38, 0 и от 1.18E-38 до 3.40E+38 Диапазон значений для двойной точности: от -1.79E+308 до -2.23E-308, 0 и от 2.23E-308 до 1.79E+308</p>
8	DOUBLE	<p>Тип с плавающей точкой размером 8 байт (двойная точностью)</p>

9	BIT	<p>Целочисленный тип данных, который может принимать значения 0, 1 или NULL (используется для хранения битовых значений)</p> <p>BIT (n), где n — количество битов (от 1 до 64), если n указано больше, чем длина хранимого значения, то значение слева заполняется нулями, например, для числа 13 — 1101, если задано BIT (7), то значение примет следующий вид — 0001101</p>
10	DATE	<p>Хранение даты в формате YYYY-MM-DD</p> <p>Поддерживает диапазон от 1000-01-01 до 9999-12-31</p>
11	DATETIME	<p>Хранение даты и времени в формате YYYY-MM-DD hh:mm:ss</p> <p>Поддерживает диапазон от 1000-01-01 00:00:00 до 9999-12-31 23:59:59</p>
12	TIME	<p>Хранение значения времени в формате hh:mm:ss</p> <p>Поддерживает диапазон от -838:59:59 до 838:59:59</p> <p>Используется не только для представления времени дня (которое должно быть меньше 24 часов), но и для прошедшего времени или временного интервала между двумя событиями</p>
13	YEAR	<p>Хранение значения года в формате YYYY</p> <p>Тип YEAR занимает 1 байт, поэтому поддерживает диапазон от 1901 до 2155 и 0000 (MySQL 8.0 не поддерживает задание года в двузначном формате)</p>

14	CHAR	Строка фиксированной длины (может содержать буквы, цифры и специальные символы). CHAR (size), где size — длина строки в символах (от 0 до 255, по умолчанию 1)																									
15	VARCHAR	Строка переменной длины (может содержать буквы, цифры и специальные символы). VARCHAR (size), где size — максимальная длина строки в символах (от 0 до 65535) Разница между CHAR и VARCHAR <table><tr><td>Значение</td><td>CHAR (4)</td><td>Байт</td><td>VARCHAR (4)</td><td>Байт</td></tr><tr><td>' '</td><td>' '</td><td>4</td><td>' '</td><td>1</td></tr><tr><td>'AB'</td><td>'AB '</td><td>4</td><td>'AB'</td><td>3</td></tr><tr><td>'ABCD'</td><td>'ABCD'</td><td>4</td><td>'ABCD'</td><td>5</td></tr><tr><td>'ABCDEFGH'</td><td>'ABCD'</td><td>4</td><td>'ABCD'</td><td>5</td></tr></table> При использовании однобайтовых кодировок размер типов CHAR и VARCHAR при хранении равен количеству символов (VARCHAR помимо самой строки еще хранит префикс длины — количество байтов строки). Однако в случае многобайтовых кодировок, таких как UTF-8, в старших диапазонах Юникода один символ занимает два или несколько байт	Значение	CHAR (4)	Байт	VARCHAR (4)	Байт	' '	' '	4	' '	1	'AB'	'AB '	4	'AB'	3	'ABCD'	'ABCD'	4	'ABCD'	5	'ABCDEFGH'	'ABCD'	4	'ABCD'	5
Значение	CHAR (4)	Байт	VARCHAR (4)	Байт																							
' '	' '	4	' '	1																							
'AB'	'AB '	4	'AB'	3																							
'ABCD'	'ABCD'	4	'ABCD'	5																							
'ABCDEFGH'	'ABCD'	4	'ABCD'	5																							
16	TINYTEXT	Хранение строки максимальной длины в 255 символов																									
17	TEXT	Хранение строки максимальной длины в 65 535 символов																									
18	MEDIUMTEXT	Хранение строки максимальной длины в 16 777 215 символов																									
19	LONGTEXT	Хранение строки максимальной длины в 4 294 967 295 символов																									

20	BINARY	Аналог CHAR, но данные хранятся в виде бинарной строки (бинарная строка состоит только из символов 0 и 1) BINARY (size), где size — длина строки в байтах (от 0 до 255, по умолчанию 1)
21	VARBINARY	Аналог VARCHAR, но данные хранятся в виде бинарной строки VARBINARY (size), где size — максимальная длина строки в байтах (от 0 до 65535)
22	TINYBLOB	Хранение Binary Large Object (двоичный большой объект) размером до 255 байт включительно BLOB предназначен для хранения изображений или скомпилированного программного кода
23	BLOB	Хранение BLOB размером до 65 535 байт включительно
24	MEDIUMBLOB	Хранение BLOB размером до 16 777 215 байт включительно
25	LONGBLOB	Хранение BLOB размером до 4 294 967 295 байт включительно
26	ENUM	Специальный строковый тип, который принимает <u>только одно значение</u> из фиксированного списка значений. В списке ENUM, который определяется <u>во время создания таблицы</u> в базе данных, можно задать до 65 535 значений. Все недопустимые значения (которых нет в списке) при добавлении заменяются на пустые строки. Значения в списке отсортированы в порядке их ввода.