

§ 27. Основные понятия

Что такое база данных

База данных (БД) — совокупность определенным образом организованной информации на какую-то тему (в рамках некоторой предметной области). Например:

- база данных книжного фонда библиотеки;
- база данных кадрового состава учреждения;
- база данных законодательных актов в области уголовного права;
- база данных современной эстрадной песни.

Конечно, вся эта информация может храниться и на бумаге (например, книжный каталог библиотеки). Но современным средством хранения и обработки баз данных является, безусловно, компьютер. В дальнейшем мы будем иметь в виду только компьютерные БД.

Базы данных бывают фактографическими и документальными.

В фактографических БД содержатся краткие сведения об описываемых объектах, представленные в строго определенном формате. Из приведенных выше примеров две первые БД скорее всего будут организованы как фактографические. В БД библиотеки о каждой книге хранятся библиографические сведения: год издания, автор, название и пр. Разумеется текст книги в ней содержаться не будет. В БД отдела кадров учреждения хранятся анкетные данные сотрудников: фамилия, имя, отчество; год и место рождения.

Базы данных в третьем и четвертом примерах наверняка будут организованы как документальные. Первая из них будет включать в себя тексты законов; вторая — тексты и ноты песен; биографическую и творческую справочную информацию о композиторах, поэтах, исполнителях; звуковые записи и видеоклипы. Следовательно, *документальная БД содержит обширную информацию самого разного типа: текстовую, графическую, звуковую, мультимедийную.*

Современные информационные технологии постепенно стирают границу между фактографическими и документальными БД. Существуют средства, позволяющие легко подключать любой документ (текстовый, графический, звуковой) к фактографической базе данных.

Сама по себе база данных не может обслужить запросы пользователя на поиск и обработку информации. БД — это только «информационный склад». Обслуживание пользователя осуществляет информационная система.

Информационная система — это совокупность базы данных и всего комплекса аппаратно-программных средств для ее хранения, изменения и поиска информации, для взаимодействия с пользователем.



Примерами информационных систем являются системы продажи билетов на пассажирские поезда и самолеты. WWW — это тоже пример глобальной информационной системы.

Далее в нашей книге речь будет идти только о фактографических базах данных. Дадим более строгое определение компьютерной БД, чем то, что приводилось выше.

База данных — организованная совокупность данных, предназначенная для длительного хранения во внешней памяти ЭВМ и постоянного применения.



Для хранения БД может использоваться как один компьютер, так и множество взаимосвязанных компьютеров.

Если различные части одной базы данных хранятся на множестве компьютеров, объединенных между собой сетью, то такая БД называется **распределенной базой данных**.

Очевидно, информацию в Интернет, объединенную паутиной WWW, можно рассматривать как распределенную базу данных. Распределенные БД создаются также и в локальных сетях.

Реляционные базы данных

Информация в базах данных может быть организована по-разному. Чаще всего используется табличный способ.

Базы данных с табличной формой организации называются **реляционными БД**.

В чем же их преимущество?

Главное достоинство таблиц — в их понятности. С табличной информацией мы имеем дело практически каждый день. Загляните, например в свой дневник: расписание занятий там представлено в виде таблицы, ведомость с оценками за четверти имеет табличный вид. Когда мы приходим на вокзал, смотрим расписание электричек. Какой вид оно имеет? Это таблица! А еще есть таблица футбольного чемпионата. И журнал учителя, куда он ставит вам оценки, — тоже таблица.

Видите, как много примеров, и их еще можно продолжить. Мы настолько привыкли к таблицам, что обычно не требуется никому объяснять, как ими пользоваться. Ну разве что маленькому ребенку, который только учится читать.

В реляционных БД строка таблицы называется *записью*, а столбец — *полем*. В общем виде это выглядит так:

	поле 1	поле 2	поле 3	поле 4	поле 5
запись 1
запись 2
...					

Табл. 6.1-6.5 будем в дальнейшем рассматривать как примеры информации, пригодной для организации реляционных баз данных.

Каждое поле таблицы имеет имя. Например, в табл. 6.2 «Погода» имена полей такие: ДАТА, ОСАДКИ, ТЕМПЕРАТУРА, ДАВЛЕНИЕ, ВЛАЖНОСТЬ.

Одна запись содержит информацию об одном объекте той реальной системы, модель которой представлена в таблице.

Например, домашняя библиотека — это множество книг. Значит, отдельный объект такой системы — это книга, и одна запись в базе данных «Домашняя библиотека» (табл. 6.1) — это информация об одной книге из библиотеки.

Поля — это различные характеристики (иногда говорят — атрибуты) объекта. Значения полей в одной строчке относятся к одному объекту.

Разные поля отличаются именами. А чем отличаются друг от друга разные записи? Записи различаются значениями ключей.

Главным ключом в базах данных называют поле (или совокупность полей), значение которого не повторяется у разных записей.

В БД «Домашняя библиотека» разные книги могут иметь одного автора, могут совпадать названия книг, год издания, полка. Но инвентарный номер у каждой книги свой (поле НОМЕР). Он-то и является главным ключом для записей в этой базе данных. Главным ключом в БД «Погода» является поле ДЕНЬ, так как его значение не повторяется в разных записях.

Не всегда удается определить одно поле в качестве ключа. Вот, например, база данных, которая хранится в компьютере управления образования области. В ней содержатся сведения о всех средних школах районных центров в виде такой таблицы:

Таблица 7.1. Школы

Город	Номер шк.	Директор	Адрес	Телефон
Крюков	1	Иванов А.П.	Пушкина, 5	12-35
Шадринск	1	Строев С.С.	Лесная, 14	4-23-11
Шадринск	2	Иванов А.П.	Мира, 34	4-33-24
.....

В такой таблице у разных записей не могут совпасть только одновременно два поля ГОРОД и НОМЕР ШКОЛЫ. Эти два поля вместе образуют составной *ключ*: ГОРОД-НОМЕР ШКОЛЫ. Составной ключ может состоять и более чем из двух полей.

С каждым полем связано еще одно очень важное свойство — *тип поля*.

Тип определяет множество значений, которые может принимать данное поле в различных записях.

В реляционных базах данных используются четыре основных типа полей:

- числовой;
- символьный;
- дата;
- логический.

Числовой тип имеют поля, значения которых могут быть только числами. Например, в БД «Погода» три поля числового типа: ТЕМПЕРАТУРА, ДАВЛЕНИЕ, ВЛАЖНОСТЬ.

Символьный тип имеют поля, в которых будут храниться символьные последовательности (слова, тексты, коды и т.п.). Примерами символьных полей являются поля АВТОР и НАЗВАНИЕ в БД «Домашняя библиотека»; поле ТЕЛЕФОН в БД «Школы».

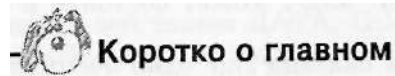
Тип «дата» имеют поля, содержащие календарные даты в форме «день/месяц/год» (в некоторых случаях используется американская форма: месяц/день/год). Тип «дата» имеет поле ДЕНЬ в БД «Погода».

Логический тип соответствует полю, которое может принимать всего два значения: «да» — «нет» или «истина» — «ложь» или (по-английски) «true» — «false». Если двоичную матрицу представить в виде реляционной БД (табл. 6.4, 6.5), то ее полям, принимающим значения «0» или «1», удобно поставить в соответствие логический тип. При этом «1» заменится на значение «истина», «0» — на значение «ложь».

Итак, значения полей — это некоторые *величины определенных типов*.

От типа величины зависят те действия, которые можно с ней производить.

Например, с числовыми величинами можно выполнять арифметические операции, а с символьными и логическими — нельзя.



База данных — организованная совокупность данных, предназначенная для длительного хранения во внешней памяти ЭВМ и постоянного применения.

Фактографическая БД содержит краткие сведения об описываемых объектах, представленные в строго определенном формате.

Документальная БД содержит обширную информацию самого разного типа: текстовую, графическую, звуковую, мультимедийную.

Распределенной называется база данных, разные части которой хранятся на различных ЭВМ компьютерной сети.

Информационная система — это совокупность базы данных и всего комплекса аппаратно-программных средств для ее хранения, изменения и поиска информации, для взаимодействия с пользователем.

Реляционные базы данных содержат прямоугольные таблицы. Строка таблицы называется записью, столбец — полем.

Таблица имеет главный ключ, отличающий записи друг от друга. Ключом может быть одно поле (простой ключ) или несколько полей (составной ключ). "

Каждое поле таблицы имеет свое уникальное имя и тип. Тип определяет, какого рода информация хранится в поле и какие действия над ней можно производить.

В БД используются четыре основных типа полей: числовой, символьный, логический, дата.

Вопросы и задания

1. Что такое база данных?
2. В чем различие между фактографическими и документальными БД?
3. Что такое распределенная БД?
4. Что такое информационная система? Приведите примеры информационных систем.
5. Что такое реляционная БД?
6. Что такое запись, поле записи? Какую информацию они содержат?
7. Определите имена полей в таблицах «Домашняя библиотека» (табл. 6.1), «Факультативы» (табл. 6.5), «Погода» (табл. 6.2), «Успеваемость» (табл. 6.3).
8. Что такое главный ключ записи? Какие бывают ключи?
9. Назовите объекты, сведения о которых содержат записи баз данных «Погода», «Факультативы», «Успеваемость». Определите ключи записей в этих БД.
10. Определите следующие понятия: имя поля, значение поля, тип поля. Какие бывают типы полей? Что обозначает каждый из типов?
11. Определить типы всех полей в таблицах «Домашняя библиотека», «Погода», «Школы».
12. Определите структуру (состав полей), ключи и типы полей для реляционных баз данных под такими названиями:
 - «Страны мира»;
 - «Мои одноклассники»;
 - «Кинофильмы»;
 - «Телефонный справочник»;
 - «Мои посещения врача».