1. **Теоретические основы БД. Сущность. Атрибут. Структуризация информации**

Итак, целью любой ИС является **обработка *данных* об *объектах*** реального мира с **учетом *связей* между ними**. В теории баз данных (БД) **данные часто называют *атрибутами*, а объекты – *сущностями***. Объект, атрибут и связь – фундаментальные понятия ИС в теории БД.

**Объект (сущность)** – нечто существующее и различимое, для которого существует название и способ отличать от другого объекта.

Например, каждая школа это объект. Объектами также являются человек, класс в школе, предприятие, химическое соединение и т.д. Объектами м.б. не только материальные предметы, но и более абстрактные понятия, Отражающие реальные мир. Например, события, регионы, произведения искусств, книги и т.д.

**Атрибут (данное)** – это определенный показатель, который характеризует объект и принимает для конкретного экземпляра этого объекта некоторое числовое, текстовое или иное значение.

Например, пусть объектом является продукция, производимой фирмой. Тогда данными (атрибутами) этой продукции будут ее код, название, цвет, вес, дата выпуска и т.д. Другой пример набора объектов – группы факультета. Число студентов в группе – это атрибут, который принимает числовое значение. Название группы – это атрибут, принимающий текстовое значение. В свою очередь, атрибут некоторого набора объектов сам м.б. объектом или набором объектов. Например, атрибутом объекта факультет являются группы студентов, которые ему принадлежат.

Т.о. возникает возможность устанавливать **связи** между объектами из разных наборов.

Информацию об объектах можно и нужно структурировать, т.к. структурированную информацию значительно легче обрабатывать.

**Структурирование информации** – это введение каких либо соглашение о способах представления данных (процесс приспособления форматов и значений данных к нуждам автомата).

Например, текстовые файлы (и документы формата Word) содержат неструктурированную или в лучшем случае плохо структурируемую информацию, непригодную для обработки автоматом.

Забегая вперед, следует сказать, что если часть информации в БД будет храниться в метрах, часть в футах, а часть и вовсе в дюймах, то обработка такой информации будет крайне не эффективна и медленна.

Пример неструктурированных данных:

Иванов Петр, Дата рождения 31.05.76, №21

№22, Андрей Сергеенко, 31/05/76

Спицын Александр Павлович, 31 мая 1972 год, №56

Список данных, организованный таким образом довольно сложно представить. Еще сложнее организовать формализованный поиск в таком списке. Чтобы автоматизировать поиск и систематизировать эти данные, необходимо выработать определенные соглашения о способах представления информации. Т.е. задать порядок представления данных, фиксировав длину полей для каждого атрибута.

Преобразования данных от произвольного вида к строгому представлению называется **структуризацией**.

Результат выполнения структуризации для рассматриваемого примера представлен в таблице 1.

|  |  |  |  |
| --- | --- | --- | --- |
| № | Фамилия | Имя | Дата свершения сделки |
| 21 | Иванов | Петр | 31/05/76 |
| 22 | Сергеенко | Андрей | 31/05/76 |
| 56 | Спицын | Александр | 31/05/72 |