

2.2 DATA DICTIONARY

- A data dictionary is a catalogue a repository of the elements in a system. These dictionary mainly talks about data and the way they are structured to meet user requirements and organizational needs.
- A data dictionary is a structured repository which contains information about the data in the system. It gives the contents of each data flow and data store in a system in a structured way. It is an alphabetic arrangement of this stored information.
- A Data Dictionary (DD) stores an organized collection of information about data and their relationships, data-flows, data types, data stores, processes and so on.
- In addition, it helps users to understand the data types and processes defined along with their uses. It also facilitates the validation of data by avoiding duplication of entries and provides the users with an online access to definitions.

Definition:

- A data dictionary refers to, "a collection of descriptions of the data objects or items in a system". **OR**
- A data dictionary can be defined as, "a centralized repository of information about system and its elements such as meaning, relationships, origin, usage and format".
- Although the format of data dictionaries varies from method to method, most use the called as project (See Fig. 2.23).
- Data element sheet and information about the data item shown in Fig. 2.23.

Name	Alias	Use	Content Description	Supplement/Additional Information

Fig. 2.23

- Various data items of DD in Fig. 2.23 are explained below:
 1. **Name:** A formal or primary name of the data or control item data storage or external entity.
 2. **Alias:** Other names or acronyms of data item.
 3. **Use:** A listing of processes which use the data or control item. And how or when it is used.
 4. **Content description:** Standard format for representing the contents of data or control item.
 5. **Supplement/Additional Information:** Other information such as limitations, restriction defaults or initial value etc.
- For example, we take an example of Integrated office phone system, (See Fig. 2.24).

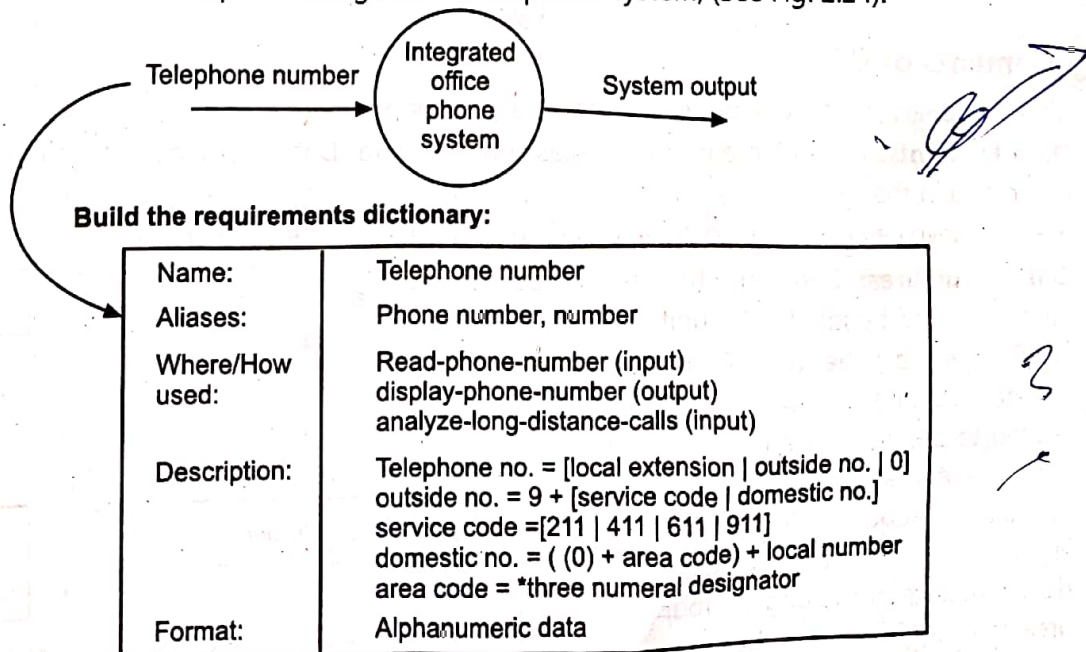


Fig. 2.24

- The description format is frequently given a highly structured form but it is to be noted that consistency in naming must be followed. Except the content description all the information may be provided directly.
- To develop the content description of a data item some notations (symbols) as shown in Fig. 2.25.

Data Construct	Notation	Meaning
	=	is composed of
Sequence	+	and
Selection	[]	either-or
Repetition	{ } ⁿ	n repetitions of
	()	optional data
	* ... *	delimits comments

Fig. 2.25

- Above notation enables representation of composite data in one of the three fundamental ways that it can be constructed:
 - As a sequence of data items.
 - As a selection from among a set of data items.
 - As a repeated grouping of data items. Each data item entry that is represented as part of a sequence, selection, or repetition may itself be another composite data item that needs further refinement within the dictionary.
- As an example using the Figs. 2.24 and 2.25 a telephone number might appear in data dictionary as shown Fig. 2.26.

Name	Alias	Use	Content Description	Supplement
Telephone number	None	Phone setup dial phone	Telephone number = local number long distance number] local number = prefix + access number long distance number = 0 + area code + local number prefix = [2520/2524] access number any three digit string area code = [141,151,]	None

Fig. 2.26

2.2.1 Elements of DD

The data dictionary has following types of components or element:

- Data Elements:** Data element is a smallest unit of data. Data elements are building blocks for all other data in the system. It is a smallest unit of data which cannot be meaningfully decomposed. For example, Invoice number, invoice date and amount due are the data elements.
- Data Structures:** Data structure is a group of data elements handled as a unit. Data elements are grouped together to makeup a data structure. A data structure is a set of data items that are related to are another and collectively describe a component in a system. Data structure shows the information about the formation of the data item. Fig. 2.27 illustrates this. For example, Phone is a data structure consisting of four data elements: area code-exchange-number-extension.

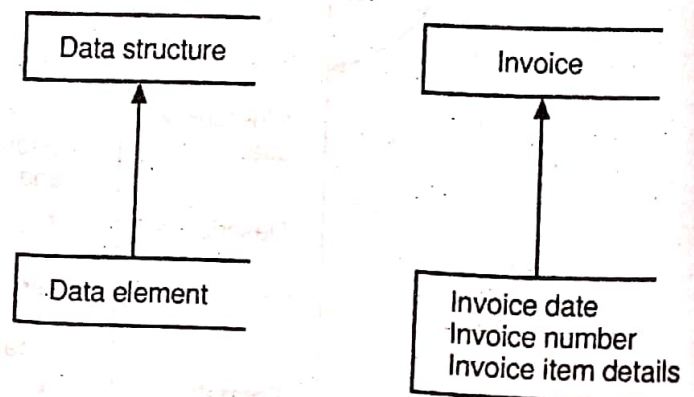


Fig. 2.27: Data Structure and Data Element

```
graph TD; A[Data Elements] --> B[Data Structures]; B --> C[Data Flow]; B --> D[Data Store];
```

Smallest Unit of Data {

Group of Data Elements {

Groups of Data Structures {

Data Elements

Data Structures

Data Flow

Data Store

The following example shows the invoice package data flow. An invoice is not approval until it has been verified against an authorized purchase order. So the invoice are delayed and they are botched for the next step which is verification.

4. **Data Names:** The names given for data elements which should be different than others like Name, Age, Sex etc.
5. **Data Description:** It shows in detail what the data item represents in the system. It is a short description of data.
6. **Aliases:** The alternate names are called as aliases. For example, SSN can be referred as SSNO or SS-NO etc.
7. **Data Type:** It shows type of data like Numeric, Float, etc.
8. **Length:** Length means number of spaces required to store the data element value. For example, stud_ID number 4 it stores the numbers from 0 to 9999.
9. **Data Value:** In some processes only specific data values as allowed, i.e. some additional prefixes are added called as data values. For example, SYBCA001; Second Year of Bachelor in Computer Application student Roll No.
SY 001; Second Year BCA

1. A DD manage the details in large systems and DD improves consistency.
2. It is used to locate omissions and errors in the system.

3. A DD is used to communicate common meaning for all system elements.
4. It is a valuable reference in any organisation.
5. It improves analyst/user communication by establishing consistent definitions of various elements terms and procedures.

Disadvantages of DD:

1. For large organization a DD is grows rapidly in size and complexity.
2. A DD is difficult to maintain manually.
3. Some users and analysts who may use other methods find it difficult to relearn their trade and start to use DDs.

2.3 ENTITY RELATIONSHIP DIAGRAM (ERD)