# Database and Database Management System

***Prepared By:***

***Ramesh Banjade***

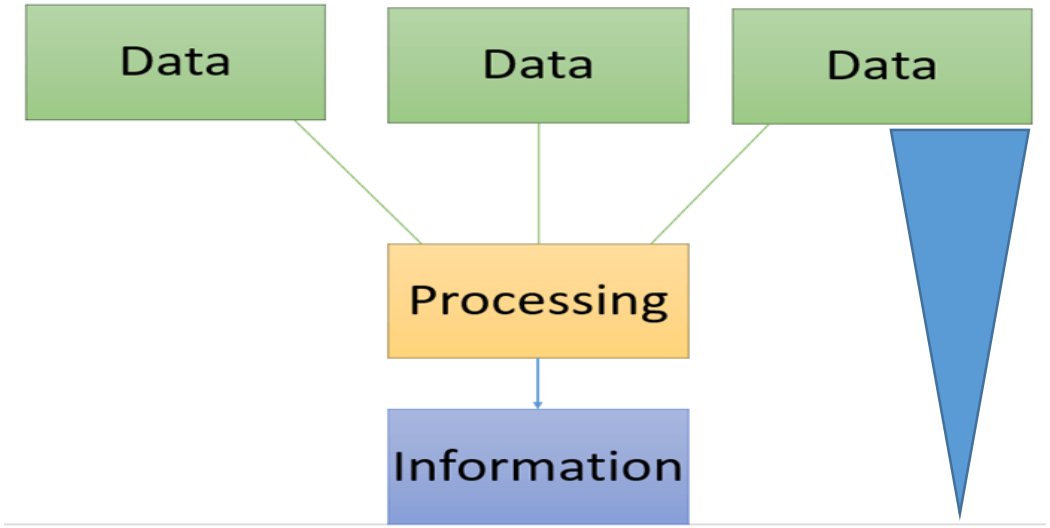
# Introduction

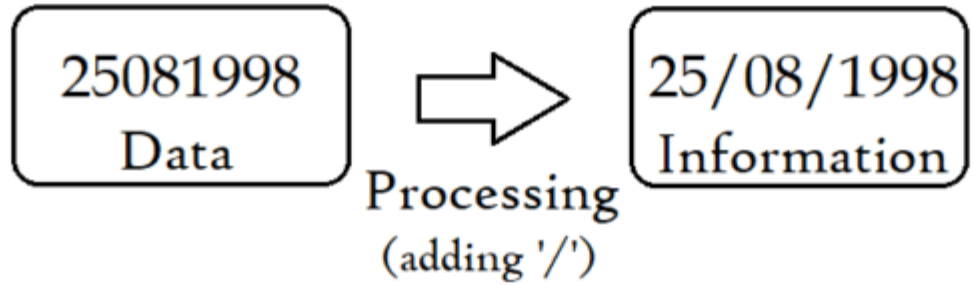
- **Data**

Data are raw fact that have no specific meaning.

- **Information**

Information is processed data that has purpose and meaning.





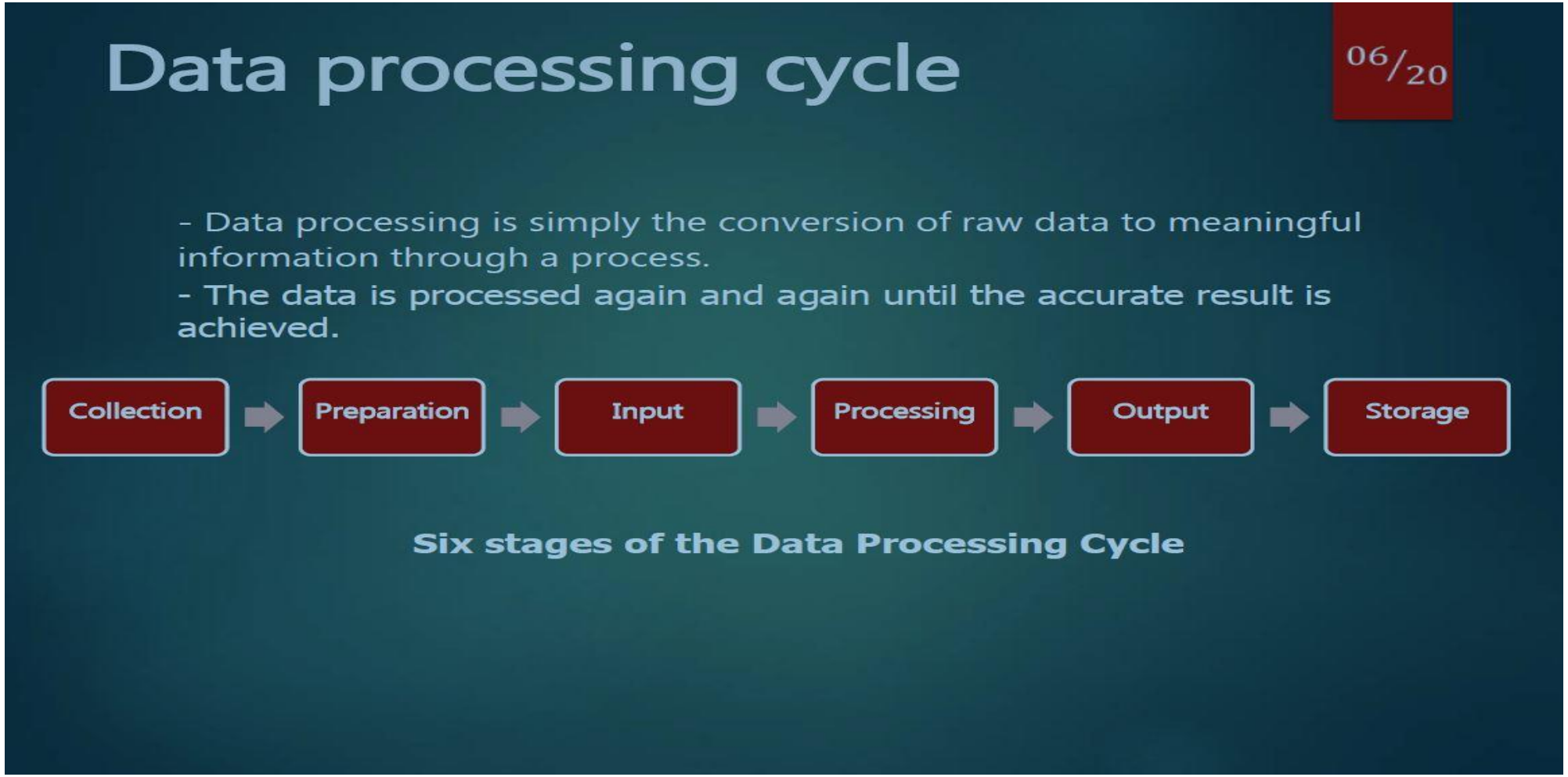

# Difference between Data and information

| Data | Information |
|---|---|
| **Data** can be any character, number, images, words, text, means little or nothing to a man. When data is processed | Organized and presented in a context to make it useful is called **information**. |
| **Data** alone can never be significant | But information is always important by itself. |
| **Data** is based on records, observations etc | **Information** is based on analysis of data |
| **Data** is unorganized and does not depends on information | But **Information** always depends on data and organized. |

# Data Processing

- Data is a collection of facts & figures, such as values or measurements. The data is the raw material to be processed by a computer.
- Data are manipulated to produce output. when this output can be used to help people make decisions, it is called information.
- **Data processing is defined as series of actions or operations that convert data into useful information.**

- Data processing also refers to the complete check of the data validity , accuracy, consistence, integrity and its relevance to the client's requirements.



**Input**

Input is the task where verified data is coded or converted into machine readable form so that it can be processed through a computer.

**Processing**

Processing is when the data is subjected to various means and

methods of manipulation, the point where a computer program is being executed, and it contains the program code and its current activity.

- classifying,
- sorting,
- calculating,
- summarizing,
- comparing

**Output**

Output is presented to users in various report formats like printed report, audio, video, or on monitor.

**Storage**

- Storage is used to store information for future use and reporting.

# Database Terminology

- **Column or  Field**
- **Row or Record**
- **Table** (similar to an Excel Worksheet)
- **Primary Key** –unique and mandatory
- **Foreign Key** –a cross-reference between tables because it references the primary key of another table
- **Relationship**–created though foreign keys.

- **One to one (1-1)**
- **One to many (1-M)**
- **Many to One (M-1)**
- **Many to Many (M-M)**

# Field, Record and Table

- **Table**

A set of data arranged in a columns and rows. The column represent characteristics/properties of stored data and the row represent the actual data entries.

| E_ID | E_NAME | E_SALARY |
|---|---|---|
| 1 | RAM | 60000 |
| 2 | SHYAM | 45000 |
| 3 | SITA | 56000 |
| 4 | GITA | 70000 |
| 5 | HARI | 80000 |

- **Field**

Fields are columns in a table with specific information about the data.

In above tale e_id, e_name and and e_salary are fields that provides the information about id , name and salary of different employees.

- **Record**

Record is a individual entry that exists in a table. It is also known as row.

Records give the complete information of a single entry or entity.

| E_ID | E_NAME | E_SALARY |
|---|---|---|
| 1 | RAM | 60000 |
| 2 | SHYAM | 45000 |
| 3 | SITA | 56000 |
| 4 | GITA | 70000 |
| 5 | HARI | 80000 |

In above table selected row gives the complete information of about the particular employee sita.

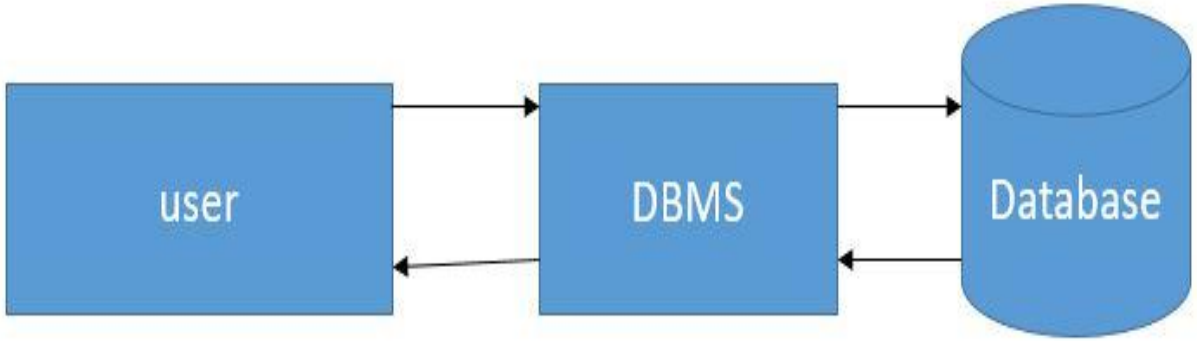
(**collections of records and fields makes a table and a collection of tables makes a database**)

# What is Database ?

- Database is an organized collection of structure data or information that is stored in such a way that it can be easily accessed , create, update, delete and managed.
- A database is usually controlled by a database Management System (DBMS)
- In Database, data are organized in row and columns in the form of table.
- Most database use structure query language (SQL) for writing and querying data.
- SQL , MYSQL, Oracle, Postgresql are the example of database.

# What is database management system?

- Database management system (DBMS) is a type of software or tools that helps to manage the database. User can create , maintain, manipulate and retrieve data or information from a single table or group of interrelated table.
- DBMS act as an interface between user/application program and database.
- Example : **Oracle Database , SQL Server, MS-access, Dbase etc. are used to managed data in DBMS.**
- DBMS application are **Student information system**, **hospital management system ,Library Management System, customer relationship Management.**



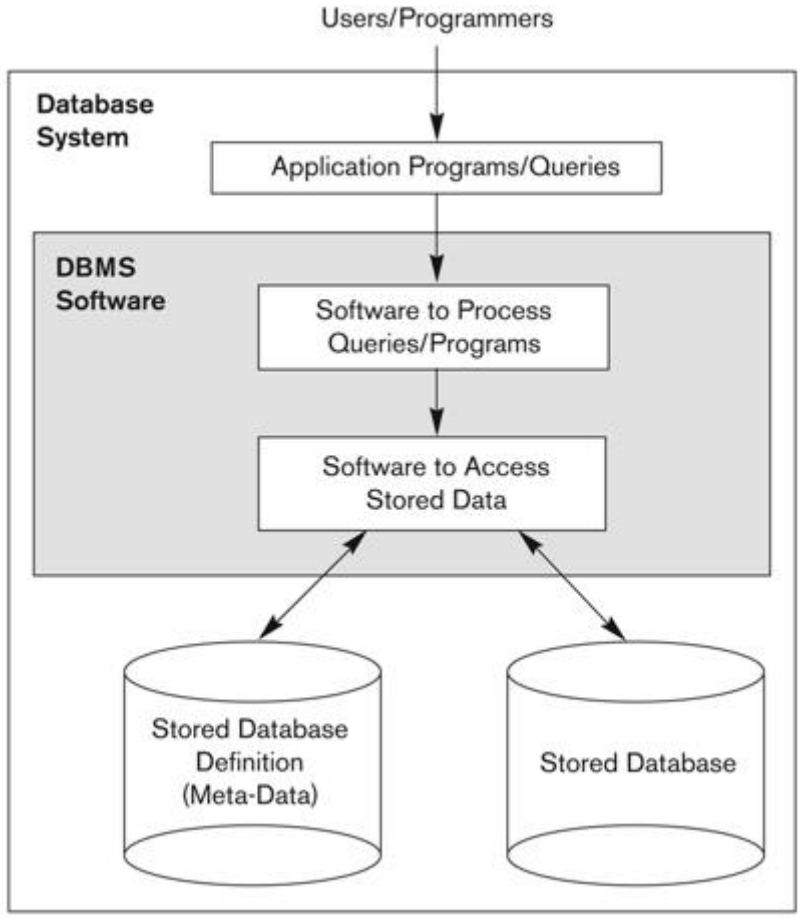




*Fig : DBMS act as interface between user and Database*

# Characteristics of the Database Approach

## 1. Manages Information

database manages all the information in their required fields (according to column definition and ids).

## 2. Easy Operation Implementation

All the operations like insert, delete, update, search, etc. by SQL queries.

Queries make it more powerful.

## 3. Multiple Views of Database

A view is a subset of the database or interface. It is defined according to user requirements. Different users of the system may have different views of the same system.

## 4. Data for Specific Purpose

- A database is designed for a specific purpose according to the group of user and application interests.
- For example, a database of the student management systems is designed to maintain the record of students' marks, fees, and attendance, etc. This data has the specific purpose of maintaining student records.

## 5. Logical Relationship Between Records and Data

- A database maintains a logical relationship between records and data. So a user can access various records according to various logical conditions by a single query from the database.

### 7. Insulation Between Program and Data

- In the database approach, the data structure is stored in the system catalog, not in the programs. If we want to change the structure of a file then no need to change the program. This feature is called program-data independence.
- It is not found in the file-based system.

### 8. Self-describing nature of a database system

- A database system is referred to as self-describing by metadata. Metadata defines and describes the data and relationships between tables in the database.
- This information is used by the DBMS software or database users if needed.

## 9. Sharing of data and multiuser system

- Current database systems can access by multiple users at the same time. This feature is called concurrency control strategies. These strategies ensure that the data accessed are always correct and integrated.

## 10. Control of data redundancy

- In the database approach, each data item is stored in only one place in the database. So single data does not repeat more than one time. This feature improves system performance.
- Redundancy is controlled by application programming and minimized when designing the database.

## 12. Restriction of unauthorized access

- A database approach should provide security to create and control different types of user accounts and restrict unauthorized access.
- It provides privileges to access or use data from a database these are read-only access (i.e., the ability to read a file but not make changes), and read and write privileges, (both read and modify a file).

## 13. Enforcement of integrity constraints

- database approach provides the ability to define and enforce certain constraints to ensure that users enter valid information and maintain data integrity.

## 13. Transaction processing

- A database approach supports a concurrency control subsystem that ensures that data remains consistent and valid during transaction processing even if several users update the same information.

## 15. Backup and recovery facilities

- Backup and recovery are methods that are used to protect your data from loss.  Backup means storing database copy to another drive or place. If a hard drive fails or is not accessible then it recovers the database from backup places.

# Major role of DBMS

- Define data structure for data storage
- Provide suitable mechanism for data access and data manipulation.
- Maintain system integrity.
- Eliminates redundancy of data
- Provide safety and security measure of data.
- Provide mechanism for data sharing among user concurrently

# Advantage of DBMS

- Make an easy to view data, add new data, modify and delete existing data.
- Organize the data in proper sequence and Structure.
- Eliminates data redundancy ( Same data set is stored in multiple places)
- Data consistency (Data consistency is the accuracy, completeness, and correctness of data stored in a database)
- Data can be shared among several users.
- Protecting data against unauthorized access.
- Allow for multiple user  to be active at one time ( Data can be accessed concurrently)
- Data Backup and Recovery.

# Disadvantage of DBMS

- Cost of Hardware and Software of a DBMS is quite high, which increases the budget of your organization.
- Most database management systems are often complex systems, so the training for users to use the DBMS is required.
- The database can fail because or power failure or the whole system stops.
- Chances of unauthorized access (security risk)
- Need more disk space for program storage then file management system.

# Objectives of the DBMS

- To  store mass storage of relevant data.
- To Make the data access easy for the user.
- To eliminate the data redundancy.
- To Make the latest modification to the database available immediately
- Allow for multiple user  to be active at one time ( Data can be accessed concurrently)
- The advancement in the database enables to add more data and program to the system
- To Protect the data from physical harm and unauthorized access.
- To Maintain Data confidentiality.

# Component of DBMS

- DBMS have several components, each performing very significant task in the database management system environment. Following are the component of DBMS and its environment.

1. Software
2. Hardware
3. Data
4. Database Access Language
5. Query processor
6. Run time database Manager/Database control system
7. Data manager
8. Database engine
9. Data dictionary
10. Report writer/ generator

1. **Software**

This is a set of program used to control and mange the overall database. This includes the DBMS software itself, the operating system itself, the network software used to share the data among users and application program used to access data in the DBMS.

2. **Hardware**

It consist of a set of physical electronic device such as computers ,I/O devices, storage devices etc. this provide the interface between computer and the real world system.

3. **Data**

The database contains both actual/operational data and the metadata.

4. **Database access language**

These is used to access the data to and form the database, to enter new data, update existing data and retrieve required data from database. the user write a set of appropriate commands in a database access language , submit these to the DBMS, which then process the data and generates and display as set of results into a user readable form. Example: SQL

**5 .Query processor**

This transforms the user queries into to a series of low level instructions. This reads the online user's query translate it into an efficient series of operations in a form capable of being sent to the run time data manager for execution.

**6. Run Time Database Manager/ Database control system**

It is the central software component of the DMBS that interfaces with user-submitted application programs and queries and handle database access at run time. It provides control to maintain the consistency, integrity and security of the data.

**7. Data manager**

This is responsible for handling of data in the database, providing a recovery to the system that allows it to recover the data after a failure.

**8. Database Engine**

The core service for storing , processing and securing data, this provides control access and rapid transaction processing to address the requirements of the most demanding data consuming application.

**9. Data dictionary**

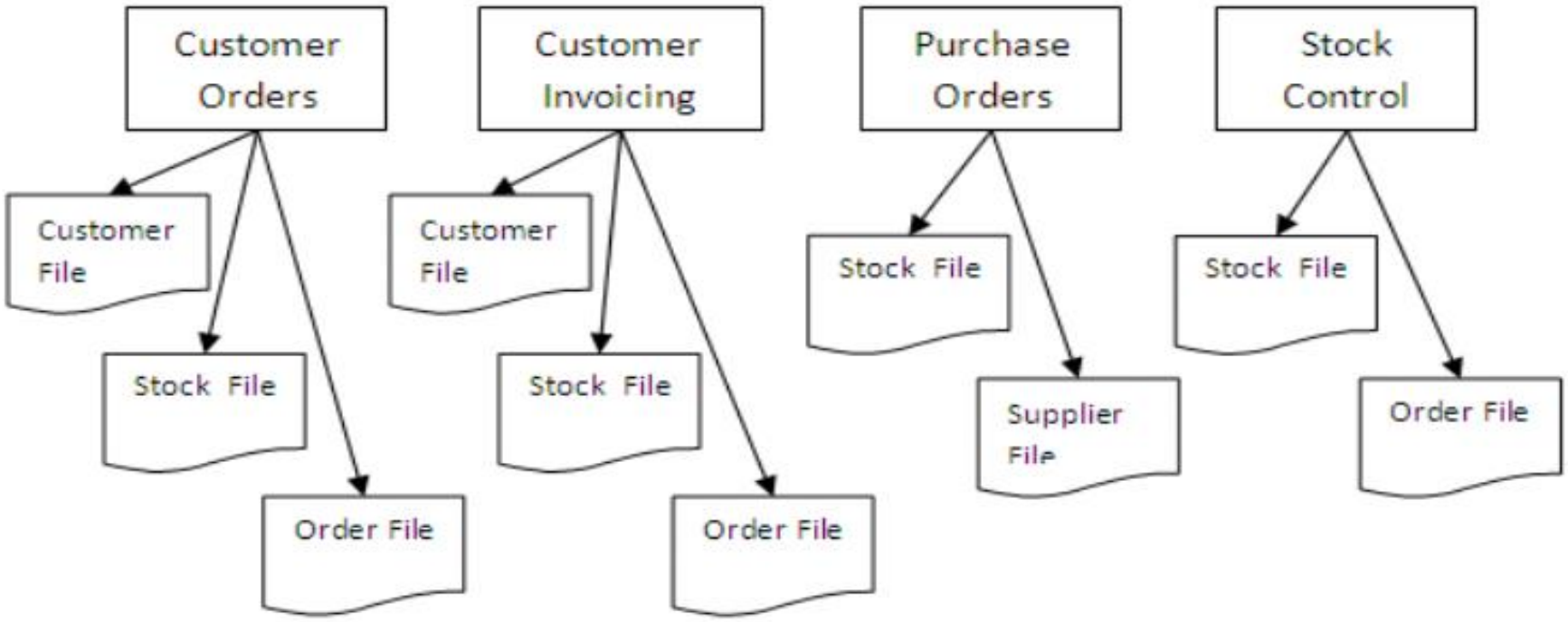
This is a reserved space within a database used to store information about the database itself. A data dictionary is a set of read only table and views, containing a different information about the data used in the enterprises to ensure that database representation of the data follow one standard as defined in the dictionary.

**10. Report writer/ Generator**

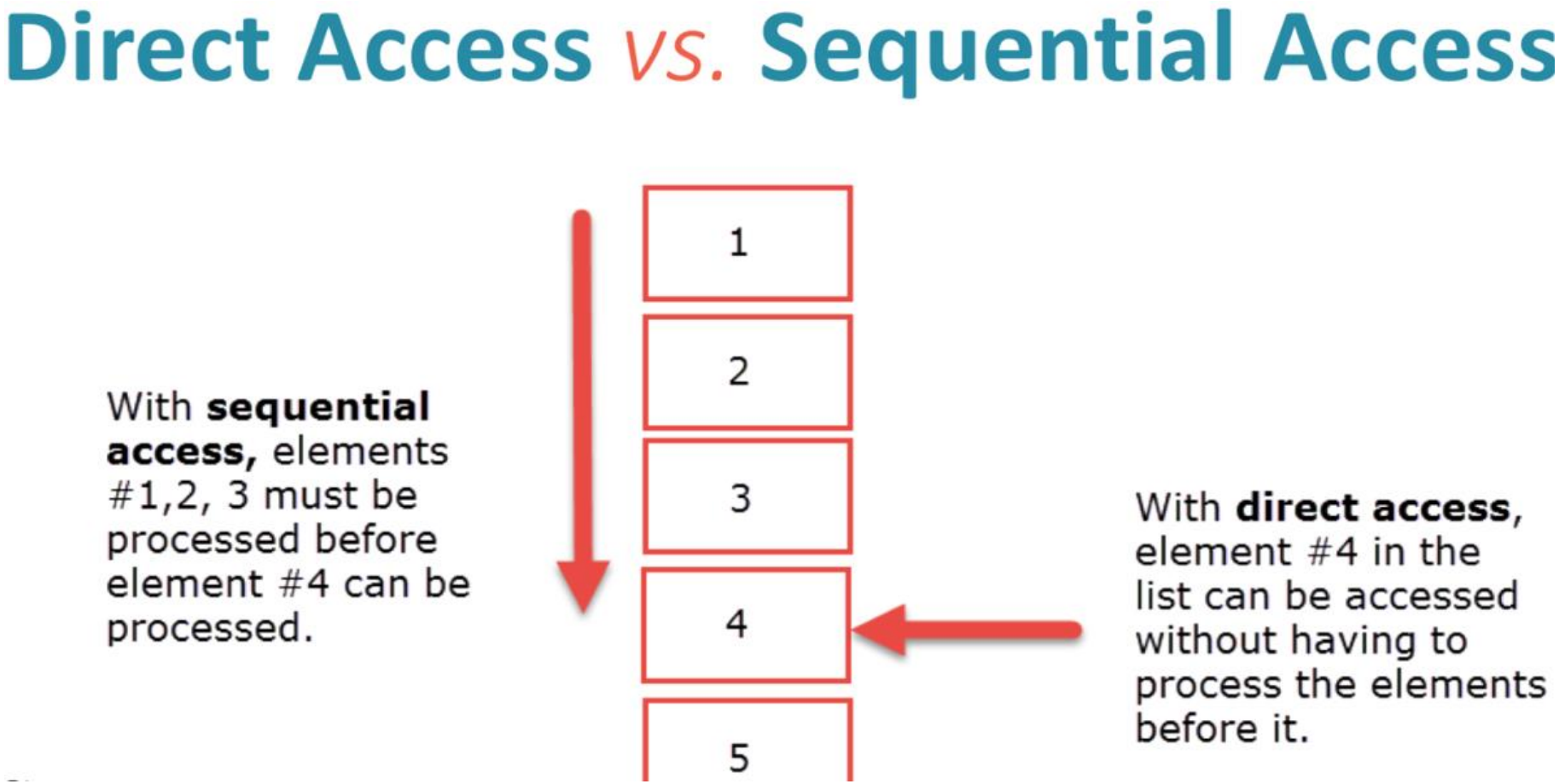
It is a program that extracts information from one or more files and present the information in a specified format .

# Traditional file-based approach

- Traditional file processing system was used to manage information. It stores data in various files of different application programs to extract or insert data to the appropriate file.



- **File processing may be either sequential or Direct access file processing.**



- This processing system has several limitation due to which database management system is required.
- **Data duplication:**

Each program stores its own separate files. If the same data is to be accessed by different programs, then each program must store its own copy of the same data.

- **Data inconsistency:**

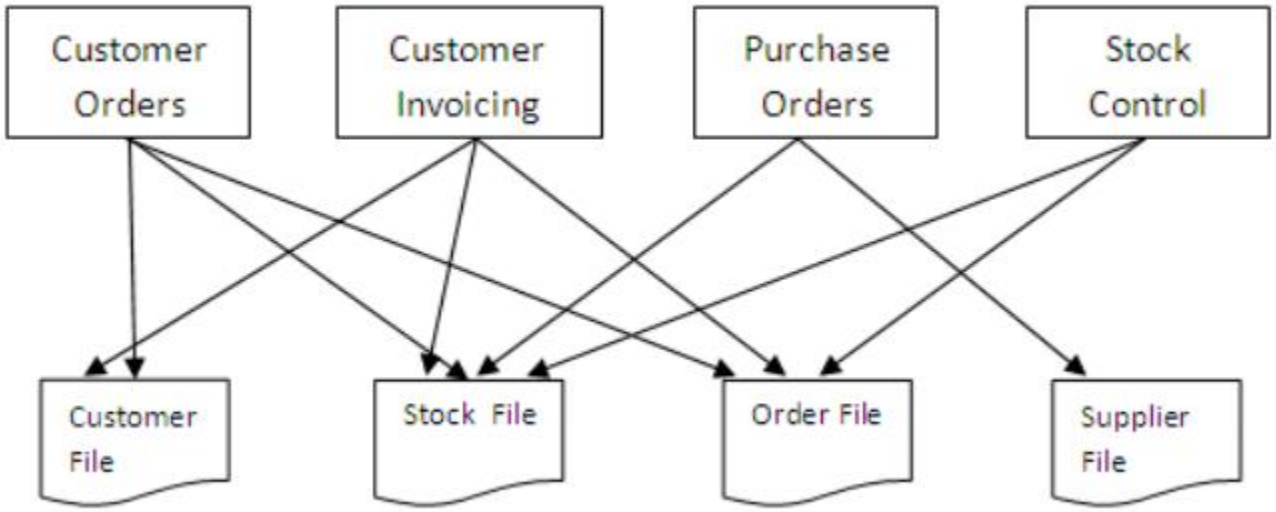
If the data is kept in different files, there could be problems when an item of data needs updating, as it will need to be updated in all the relevant files; if this is not done, the data will be inconsistent, and this could lead to errors.

- **Difficult to implement data security:**

Data is stored in different files by different application programs. This makes it difficult and expensive to implement organization-wide security procedures on the data.

## The shared file approach

- One approach to solving the problem of each application having its own set of files is to share files between different applications.




- shared files solves the problem of **duplication and inconsistent data** across different versions of the same file held by different departments, but other problems may emerge, including

- **File incompatibility:**

When each department had its own version of a file for processing, each department could ensure that the structure of the file suited their specific application. If departments have to share files, the file structure that suits one department might not suit another.

- **Difficult to control access:**

Some applications may require access to more data than others; for instance, a credit control application will need access to customer credit limit information, whereas a delivery note printing application will only need access to customer name and address details. The file will still need to contain the additional information to support the application that requires it.

- **Physical data dependence:**

If the structure of the data file needs to be changed in some way , this alteration will need to be reflected in all application programs that use that data file. This problem is known as physical data dependence.

- **Concurrent access problem**

While a data file is being processed by one application, the file will not be available for other.

# Limitation of Traditional File Processing System

- **Data duplication**
- **Data inconsistency**
- **Difficult to implement data security**
- **File incompatibility**
- **Difficult to control access**
- **Physical data dependence**
- **Concurrent access problem**

# Database system Vs. Traditional File Processing System

| Database Approach | File Based Approach |
|---|---|
| DBMS or Database Management System is a software application. It is used for accessing, creating, and managing databases. | A file system is a software that manages and organizes the files in a storage medium. It controls how data is stored and retrieved. |
| Programs and data in the database are independent. | Programs and data in the file system are interdependent. |
| Data can be retrieved and managed easily as it is stored in tables that are linked together. | Data can't be retrieved and managed easily as it is stored separately in various files. |
| Data inconsistency is higher in the file system. | Data inconsistency is low in a database management system. |
| DBMS offers good protection mechanism. | Protecting a file system is very difficult. |
| DBMS provides a crash recovery mechanism | The file system doesn't have a crash recovery mechanism. |
| DBMS system provides a concurrency facility. | These system doesn't support concurrency. |
| The redundancy of data is low in the DBMS system. | The redundancy of data is greater. |
| We can easily query data in a database using the SQL language. | There is no efficient query processing in the file system. |
| DBMS system provides backup and recovery of data even if it is lost. | It doesn't offer backup and recovery of data if it is lost |

END !!!