Introduction to Database

Instructor : Saba Iqbal

Data

- A collection of raw facts and figures related to an object
  - Object may be a person, an organization, an event, or any other thing that is significant in a system
- Data may be in form of text, numbers, images, sounds, and videos
- Collected for different purposes
- Processed to produce meaningful information
  - reports, charts, and web pages, etc.

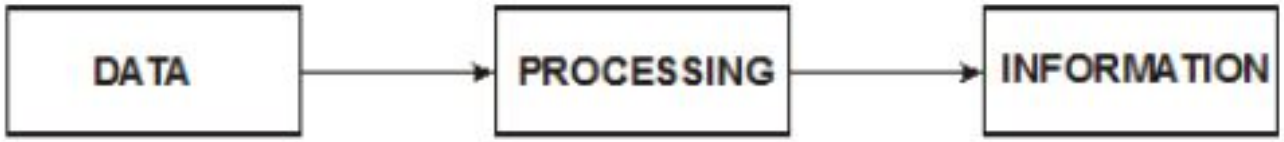
Example

- May be collected to prepare the result of an examination of students
- To completely understand it must be processed according to requirements in a system

Roll No	Name	English	Chemistry	Computer
1	Saleem	62	63	64
2	Babar	50	75	70
3	Amanat	90	80	70
4	Salma	75	80	60
--	-------	---	---	---

Information

- Usually, collected data is not in proper format and does not give proper meanings
- Various operations are performed to get the required results that give proper and useful meanings known as information
- Information means processed data
  - Then used for decision-making, analytical purposes, and fault diagnosis, etc.
- Data is processed using various techniques



Manual File-Based System

- Oldest system used for records keeping in an organization
- A set of books (or files) is prepared that contains a particular set of information
- Each department of the organization has a separate file (or set of files) for every significant task
- Information is shared among different departments through files
- Many such files are labeled and placed in one or more cabinets
- For Security, cabinets are locked or may be located in secure areas of the building

Manual File-Based System

- **Example:**
- College/University manual file-based system for maintaining records of students and courses of different faculties etc.
  - Students record file used to store records of each student
  - Students fee file used to store fee records of students
  - Students result file used to store the result of examinations of students
  - Students course file used to store data of courses taken by students

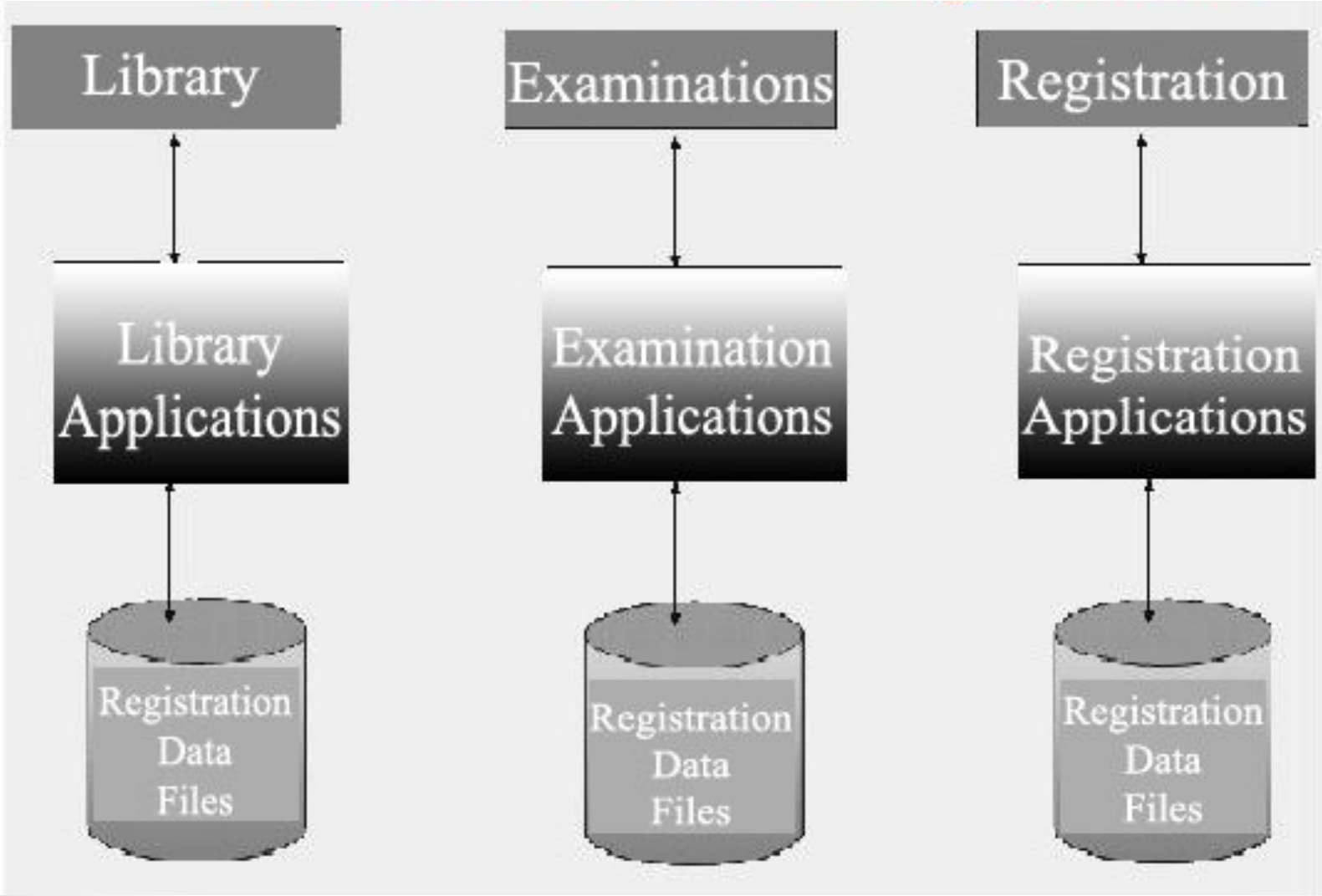
Disadvantages

- **Decentralized approach is adopted**
  - where each department processes and controls its own data
- **Very slow method to process data**
  - because data is analyzed and computed manually
  - data is transferred from one department to another manually
- **Very costly method**
  - because large number of employees have to be appointed to operate the file-based system
- **Difficulty in taking decisions**
- **Large space is needed to store the files/books**
  - Also, difficult to search a specific file or a piece of information
- **Duplication of data may exist throughout the organization**

File Processing System

- It is also known as *Computer file processing system* or *traditional file processing system* or *computer file-based system*
- In early days of computer when their is no databases
  - Data was stored in computer files on tape or disk
  - Data was stored and managed through application programs
- Still used in some small organizations
- Each department has its own set of data files
- Each application program is developed with its own set of data files that is likely to be a subset of the master file

File Processing System



Disadvantages of File Processing System

These subsets of the master file lead to

- **Data redundancy**
  - Each application has its own data file so, same data may have to be recorded and stored in many times.
- **Data inconsistency**
  - Due to the same data items that appear in more than one file do not get updated simultaneously in each and every file.
- **Data dependence**
  - Program and application in the file processing system are data dependent but, the problem is incompatible with file format.
- **Limited data sharing.**
- **The problem with security.**

Disadvantages of File Processing System

- Retrieval (retrieval is not easy).
- Time-consuming.
- Inefficient to maintain the record of the big firm having a large number of items.
- Required Lots of labor work to do.

Database

- *"Database is an organized collection of related data stored in an efficient and compact manner"*.
  - **"organized"** means that data is stored in such a way that it can easily be accessed and updated
  - **"related data"** means data and information about a particular area is stored such as:
    - Database of employees contains data of employees of that particular organization or department
    - Database of students contains data of students of a college/university etc.
  - **"efficient"** means required data can be searched very easily and quickly
  - **"compact"** means that stored data takes up as little space as possible without any duplication of data

Examples of Databases

- NADRA
- Library
- College/University
- Bank Accounts
- E-mail Accounts

Database Management System (DBMS)

- A collection of programs that are used to create, maintain, and extract data from databases
- It is general-purpose software
- This is often called database software
- Different types of DBMSs are available, ranging from small systems that run on personal computers to huge systems that run on mainframes

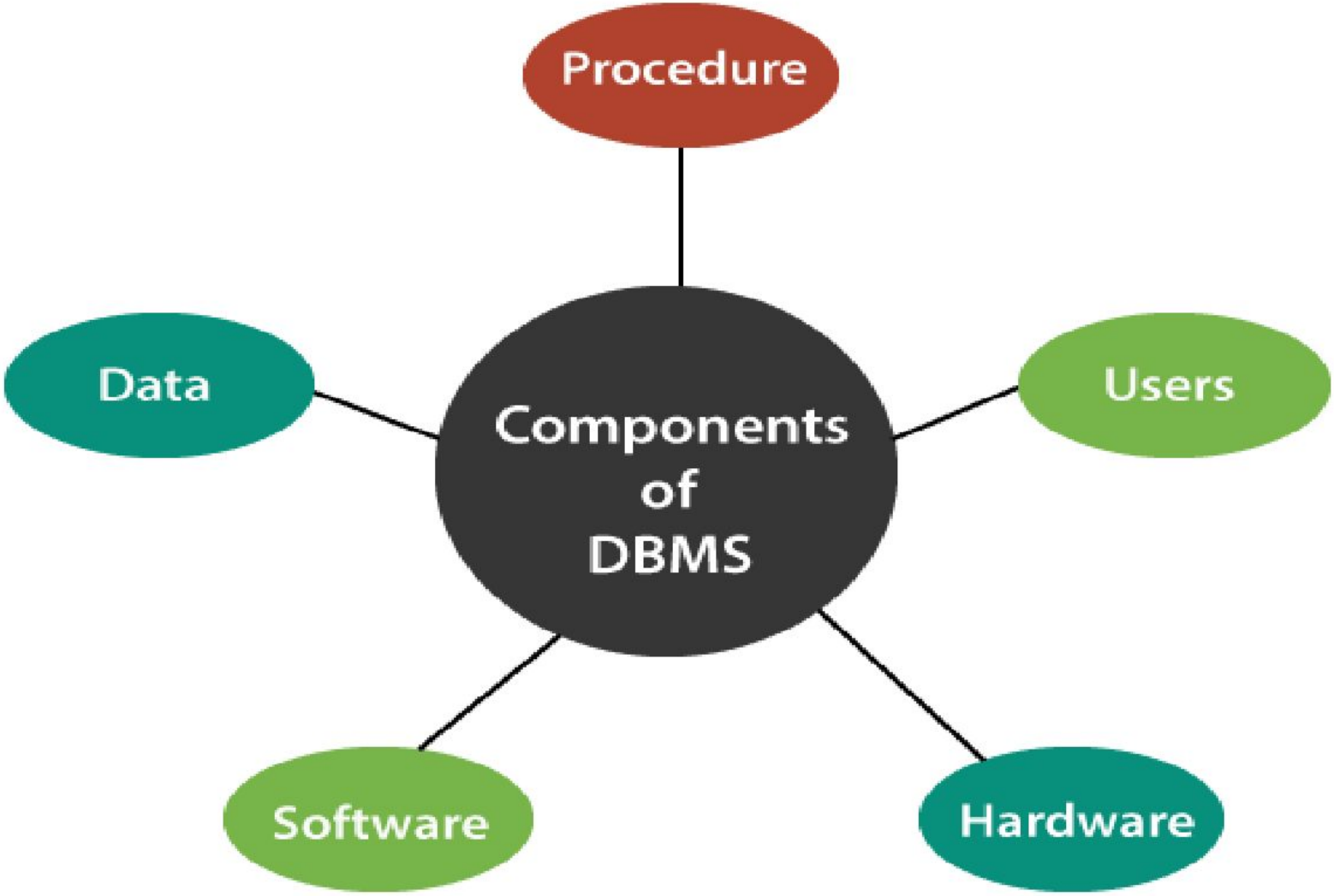
Functions of DBMS

- **Defining the Structure of Database**
  - It involves defining tables, fields and their data types, constraints for data to be stored in the database, and the relationships among tables.
- **Populating the Database**
  - Means storing data into database
- **Manipulating the Database**
  - In involves to retrieve specific data, update data, delete data, insert new data, and generate reports.

Examples of DBMS

- Microsoft Access
- Oracle
- Microsoft SQL Server
- MySQL
- FileMaker Pro

Components of a DBMS Environment



Advantages of Database

- **Controlling Data Redundancy**
  - In traditional file system, each application program has its own data files and this result in duplication of data in more than one places.
  - Database approach reduces data redundancy.
  - The data appears in a database appears only once and it is not duplicated.
  - Through controlling data redundancy storage space is saved.
- **Data Consistency**
  - By controlling data redundancy, data consistency in obtained.
  - Data item is appeared only once
  - Any update in the value is performed only once and it is readily available to everyone.

Advantages of Database

- **Data Sharing**
  - Allows to share (access) data by any number of users simultaneously.
- **Data Integration**
  - In traditional file system, data is stored in separate files and it is difficult to extract the required information.
  - In database, data is stored in tables.
  - A single database may contain multiple tables.
  - The relationships can be created between tables .
  - This makes easier to retrieve and update data.

Advantages of Database

- **Data Integrity**
  - Data integrity refers to the correctness and consistency of data.
  - It is expressed in terms of certain constraints.
  - The consistency rules are applied on database to check that the correct data is entered in the database, before storing.
- **Data Atomicity**
  - Atomicity means that either one transaction should take place as a whole or it should not take place at all.
  - If the any process is not completed successfully then the system fails, and this is known as data atomicity problem.
  - In database partially completed tasks are rolled back.
  - Only consistent data exists within the database.

Advantages of Database

- **Data Security**
  - It is the protection of the database from unauthorized users.
  - Only authorized persons are allowed to access the database.
- **Control Over Concurrency**
  - In some situations, two or more users may access the same file simultaneously.
  - It is possible that they will interfere with each other.
  - For example, two or more users are trying to update the same record, then one may overwrite the values recorded by another user. This may result in loss of information.
  - Most databases control the concurrency so that transactions are recorded with accuracy.

Advantages of Database

- **Data Independence**
  - The data stored in a database is independent of the application programs that accesses the data from database.
  - The data structure of database and application program that uses data are separate from each other.
    - The user can easily change the structure of database without modifying the application program.
    - Similarly, user can modify application programs without changing structure of database.

Disadvantages of Database

- Required large size of memory.
- Required a processor with the high speed of data processing.
- Cost of data conversion.
- Database failure (DB corrupted due to power failure or whole system stops).
- Cost of Staff training.
- Technical Staff

Lecture End