**Database -** collection of related data items within a specific business process or problem setting

A database represents some aspect of the real world, sometimes called aminiworld or the universe of discourse (UoD)

**DBMS** - a general-purpose software system that facilitates the processes of *defining, constructing, manipulating,* and *sharing*

**Defining -** involves specifying the data types, structures, and constraints of the data to be stored in the database.

**Constructing** - process of storing the data on some storage medium that is controlled by the DBMS

**Manipulating -** includes functions such as

**querying** the database to retrieve specific data

*Query*: “Which department pays the highest salary?”

**updating** the database to reflect changes in the miniworld

Increase the salaries of employees in Accounts department by 10%

**generating** reports from the data

*Report* - List monthly salaries of employees, organised by department, with average salary and total sum of salaries for each department

**Sharing -** allows multiple users and programs to access the database simultaneously

Database system

Collection of interrelated data (Database)

and a Set of programs that allow users to

access and modify these data (DBMS)

Organizations stored data in **file-processing system** before DBMSs were introduced

To allow users to manipulate the information, the system has a number of **application programs** that manipulate the files.

Data redundancy – duplicate data stored everywhere

Data inconsistency – due to redundancy there may be 2 records of the same data

Difficulty in accessing data- Need new application programs bad

Data isolation – Many programers will write different formats of data storage while writing new app prog to retrieve data the programmer strulggles as he dosent know what format is actually present

Inegrity Problems - Hard to add new constraints or change existing one

Atomicity problems - In many applications, it is crucial that, if a **failure occurs**, the data be restored to the consistent state that existed prior to the failure.

The fund transfer must be ***atomic -*** it must happen entirety or not at all

**Concurrent-access anomalies**

Concurrent access needed for performance

Main characteristics that distinguish DB approach from the file-processing approach

**Self-describing nature of a database system**

**Insulation between programs and data, and data abstraction**

**Support of multiple views of the data**

**Sharing of data and multiuser transaction processing**

**Self-Describing Nature of a Database System**

Database system contains not only the database itself but also a complete **definition or description of the database structure and constraints**

**Program data Independence – change in data no need to change in program**

**Program operation independence – change in operation no need to change program**

**Data abstraction – the characteristic which allows above 2 .**

**also hides application program to the users**

**Support of Multiple Views of the Data**

Each user may see a different view of the database, which describes **only** the data of interest to that user

**View**

Subset of the database

Contains **virtual data** derived from the database files but is not explicitly stored

**Sharing of Data and Multiuser Transaction Processing**

Allow multiple users to access the database at the same time

**Restricting Unauthorized Access is tackled with multiple views**

**Enforcing Integrity Constraints - Simplest type of integrity constraint - specifying a data type for each data** item

**Database Administrator (DBA)**

* **gives access**
* **maintaining the db**
* **Responsible for damages**

**Database designers**

* **Understand what are the requirements and design the database**

**End users**

* **Casual**
  + **Occasionally access the database**
  + **middle- or high-level managers**
* **Naive**
  + **constantly accessing the database**
  + **Bank tellers, reservation agents**
* **Sophisticated**
  + **include engineers, scientists, business analysts, and others** who thoroughly familiarize themselves with the facilities of the DBMS in order to implement their own applications to meet their complex requirements
* **Standalone**
  + **maintain personal databases by using ready-made program packages that** provide easy-to-use menu-based or graphics-based interfaces

**System Analysts and Application Programmers – System analysts will see requirements of users and make specifications and application programmer will implement this specifications as programs**

**DBMS system designers and implementers**

Design and implement the DBMS modules and interfaces as a software package

**Tool developers**

Design and implement tools - the software packages that facilitate database modeling and design, database system design, and improved performance. Tools are optional packages that are often purchased separately

**Operators and maintenance personnel (system administration** personnel)

Responsible for the actual running and maintenance of the hardware and software environment for the database system

**Data Model**

a collection of concepts that can be used to describe the **structure of a database**

**Data models are shit I cant understand anything**

**Database Schema**

**Description of a database (**overall design of the database) - **specified during database design**and is not expected to change frequently

**---- tells the basic structure of the database using the meta data**

**Database state** - the actual data stored in a database at a particular moment in time

**Also called the database instance (**or occurrence or snapshot)

**Schema** is also called intension whereas **State** is also called extension

**Valid state** - a state that satisfies the structure and constraints specified in the schema

**hree of the four important characteristics of the database approach**

(1) use of a catalog to store the database description (schema) so as to make it self-describing

(2) insulation of programs and data (program-data and program-operation independence) and

(3) support of multiple user views

**Three-schema architecture** - proposed to help achieve these characteristics

Goal of three-schema architecture - to separate the user applications from the physical database

**Internal Level – Tells us how physical data storage structures and access paths**

**Conceptual level – it will give us all the attributes that we use**

**-gives us whole structure of database**

External or view level – Shows us part of the database which the user requires / requested

**Data Independence**

the capacity to change the schema at one level of a database system without having to change the schema at the next higher level

Logical Data Independence - The capacity to change the conceptual schema without having to change the external schemas and their associated application programs

Physical Data Independence - The capacity to change the internal schema without having to change the conceptual schemas and their associated application programs