

CHAPTER ONE

Outline

- 🖱 Introduction to Database Systems
- 🖱 Data handling approach
 - 🖱 Manual approach
 - 🖱 File based approach
 - 🖱 Database approach
- 🖱 Actors on the databases
- 🖱 Database Development Life-cycle

Introduction

□ **Data are unorganized raw facts that are stored.**

- ▶ Raw facts are seldom immediately useful to a decision maker.
 - ▶ Raw facts themselves are useless
- ▶ **Data** in themselves are fairly useless. But when these data are *interpreted* and processed to determine its true *meaning*, they become useful and can be called **Information**.

Cont...

- A major requirement of any **automation** is to **store & retrieve** data in a way that is meaningful to the end user: so the overall purpose is to transform data into Information
 - Data is raw material & plain facts that need to be processed, structured or presented in a given context so as to make them useful (which is Information).
 - People often miss the **difference between data and information** and use the words interchangeably.

Cont...

- ▮ **Database:** is a collection of logically related data which designed to meet the information needs of an organization.
- ▮ **Database system:** is a system which has the elements like the data in the database, Hardware(client and server computers), Software(Application program and DBMS:DDL,DML etc.) and users(end users, application programmers, database administrator)
- ▮ **Database management system(DBMS):** is a software package/ system to facilitate the creation and maintenance of a computerized database.

Data handling Approaches

□ Data management passes through the different levels of development

👉 **Manual Approach:** in this approach, data storage and retrieval follows the primitive and traditional way of data and information handling where cards and paper are used for the purpose.

👉 There is intensive human labor.

👉 Events and objects are written on papers and cards that are kept in files.

Cont...

- ✎ Each of the files containing various kinds of information which is labeled and stored in one or more cabinet.
- ✎ The cabinet could be kept on safe place for security purpose based on the sensitivity of the information contained in it.
- ✎ Insertion and retrieval of data is done by searching first for the right cabinet then the right file and then the information
- ✎ One could have an indexing system to facilitate accessing of data

Cont...

□ **Example:** Let us build our address book

<u>Full Name</u>	<u>Telephone</u>
Abebe A.	911 5647..
Seble K.	911 5508..

□ **Limitation of Manual Approach**

- Prone to error
- Difficult to update, retrieval and integrate
- You have the data but it is difficult to compile the information
- Significant amount of duplication of data
- Cross reference is difficult

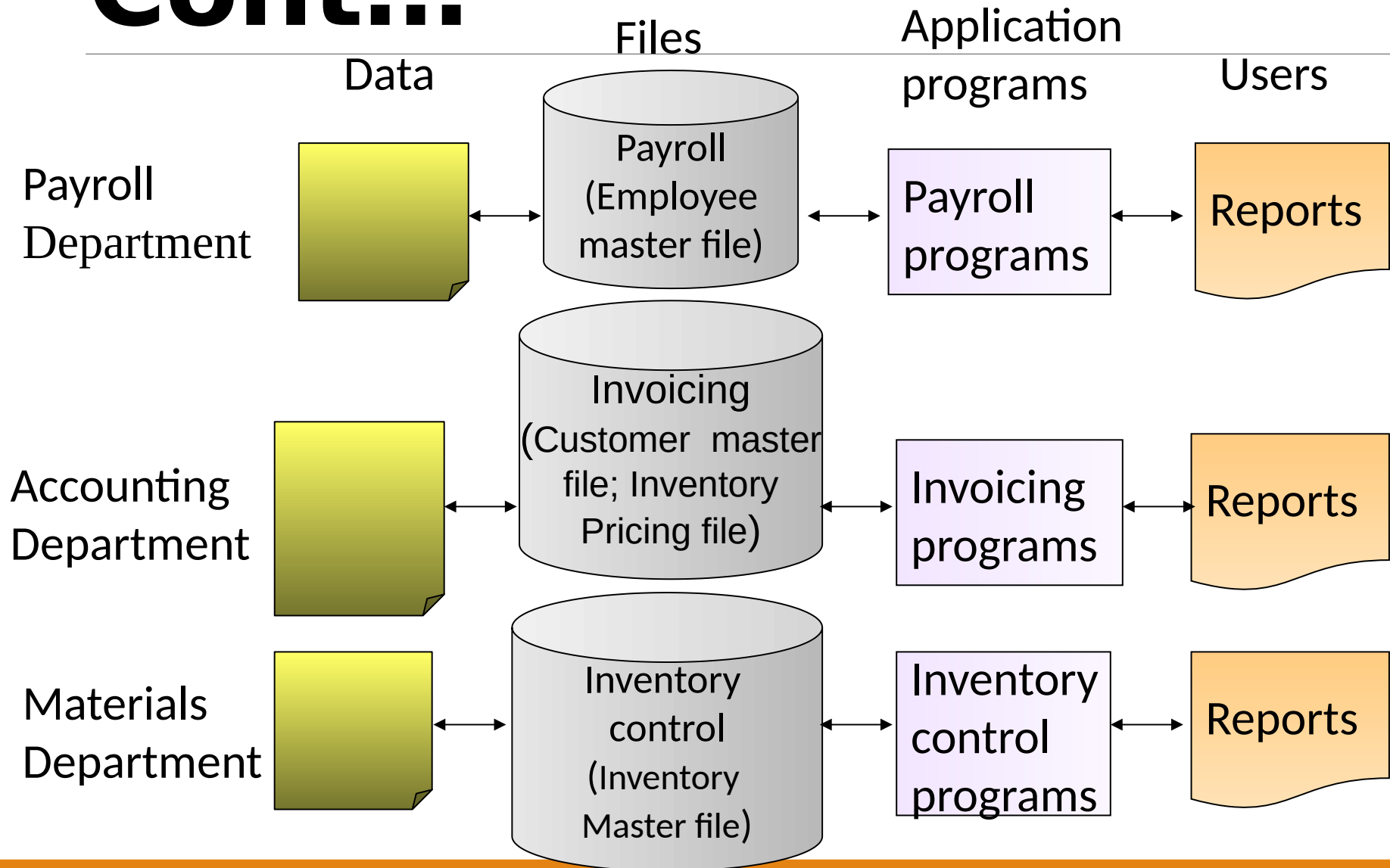
Cont...

❖ ***Solution:*** a computerized approach evolved to overcome this limitations, which is called **File based approach**.

□ **File-Based Approach:**

- It is an early attempt to computerize the manual system.
- This approach is a decentralized data handling method
 - Separate files are created and stored for each application program.
 - Develop a file or a number of files for each different application.
- Since every unit defines and manages its **own data**, the system is subjected to serious **data duplication** and **inconsistency problems**.

Cont...



Cont...

□ Limitations of File-Based systems

- ❖ **Limited data sharing:** each unit maintains its own set of data. Users of one unit may be unaware of potentially useful data held by other units.
- ❖ **Data Redundancy** (Duplication of data in different files): Same data is held by different units and hence waste storage space (Uncontrolled duplication of data)
 - ❖ **Staff_salary** (staffno, name, sex, salary)
 - ❖ **Staff_profile** (staffno, name, position, sex, DOB, salary)
- ❖ **Use of different formats for data organization:** A situation in which data organized for one application are incompatible with data organized differently for another application

Cont...

- ❖ **Separation or isolation of data:** available information in one application may not be known by other application.
- ❖ **Data dependency** on the application.
- ❖ **Data Inconsistency and confusion:** The most significant problems experienced by the traditional file based approach of data handling is the **update anomalies**. There are three types of **update anomalies**.
 - **Modification Anomalies:** A problem experienced when one or more data value is modified on one application program but not on others containing the same data set.

Cont...

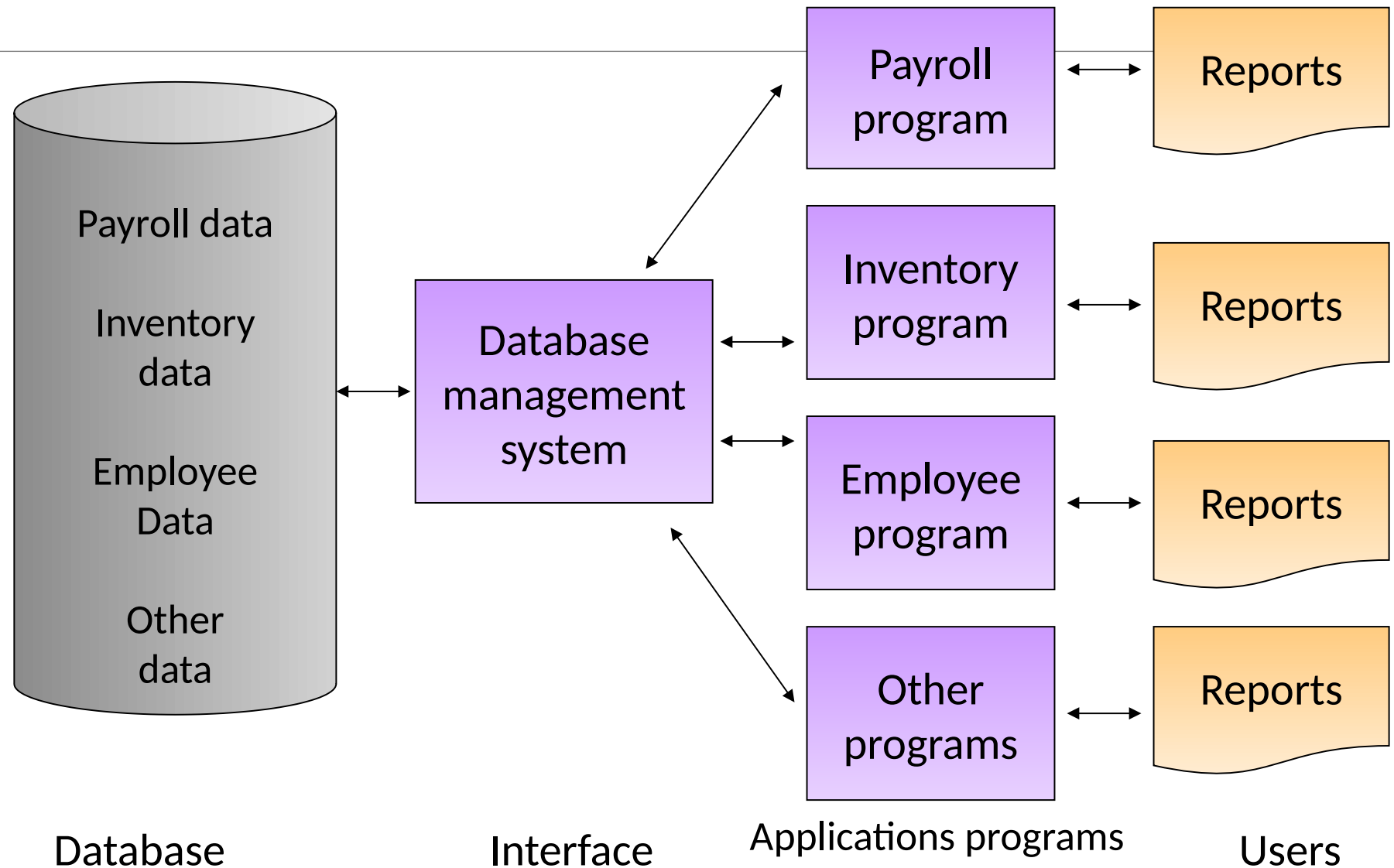
- **Deletion Anomalies:** A problem encountered where one record set is deleted from one application but remain untouched in other application programs
- **Insertion Anomalies:** A problem experienced whenever there is a new data item to be recorded, and the recording is not made in all the applications

Cont...

□ Database Approach

- To overcome the problems of file based approach, a data base approach is more effective.
- This approach is a centralized computerized data handling method
- In this approach a single repository of data is maintained.
- A pool of related data is shared by multiple applications
- Rather than having separate data files, each application uses a collection of data that is either joined or related in the database

Cont... Data base approach diagram



Cont...

▮ Advantages of data base approach

❖ **Avoid Data Duplication**

- Control data redundancy in the database, instead of storing data in redundant manner for each user.

❖ **Maintain Data Integrity**

- Data at different applications will be integrated together with additional constraints to facilitate shared data resource.

❖ **Improve decision support**

- Improved strategic use of corporate data
- The database will provide information useful for decision making

❖ **Enforce Standards**

- The different ways of using & dealing with the data by different units of the organization can be balanced & standardized by using database approach.

Cont...

❖ **Enhance Data Sharing**

- Standardization of data access. Two or more users can access and use same data/ information

❖ **Less resource (Labor, time & other) requirement**

- Unlike the other data handling methods, data maintenance will not demand much resource

❖ **Enforce Security measures**

- Better overall protection of the data
- The shared data can be secured by having different levels of clearance and other data security mechanisms.

❖ **Centralized information control**

- Since the data in the company is stored at one repository, it can be controlled & managed at the central level.
- Easier modification and updating

Cont...

▮ Disadvantages of data based approach

- ❖ Over head cost of using a data base management system
 - ▮ High initial investment in **hard ware, software, and training.**
 - ▮ Over head for providing **security, concurrency control, recovery** and **integrity** functions.
- ❖ Complexity in designing and managing data
- ❖ Reduced performance due to centralization and data independency
- ❖ High impact on the system when failure occurs to the central system.

Data base system

□ **A database** is a **very large, integrated** and **shared** collection of **logically related data** that is organized so that it can easily be **managed** (stored, searched, accessed & updated) to meet organizational information needs.

❖ Logically related data comprises entities (e.g., students, courses), attributes (ID, Name), & relationships (e.g., Student selam takes the course fundamental database system).

Cont...

- **Metadata:** the database contains not only the database itself but also a complete description of the data stored in the database.
 - This definition is stored in the system catalog, which contains information such as the structure of each file, type & storage format of each data item.
 - The information stored is called metadata since it describes the structure of the primary database.
- ❖ A **table** is the primary unit of physical storage for data in a database.
- ❖ Usually a database contains multiple tables, that are interconnected.

Cont...

- ❖ **Examples of data bases:**
 - ❖ Student records system
 - ❖ Medical records system
 - ❖ Airline bookings
 - ❖ Bank accounts
 - ❖ Stock control
 - ❖ Library catalogues
 - ❖ Product catalogues
 - ❖ Telephone directories
 - ❖ Equipment specification

Cont...

What is a Database System?

- **Database systems** are designed to organize & manage large data set in an organization and allow database users to:
 - ▶ Store their data
 - ▶ Update and maintain up-to-date data
 - ▶ Retrieve relevant data efficiently and effectively
 - ▶ Protect the database for reliable storage of the data
 - ▶ Maintain consistency of the data
 - ▶ Generate useful information for decision making
- A database system consists of data (the database), Software, Hardware, Guidelines and Users.

Cont...

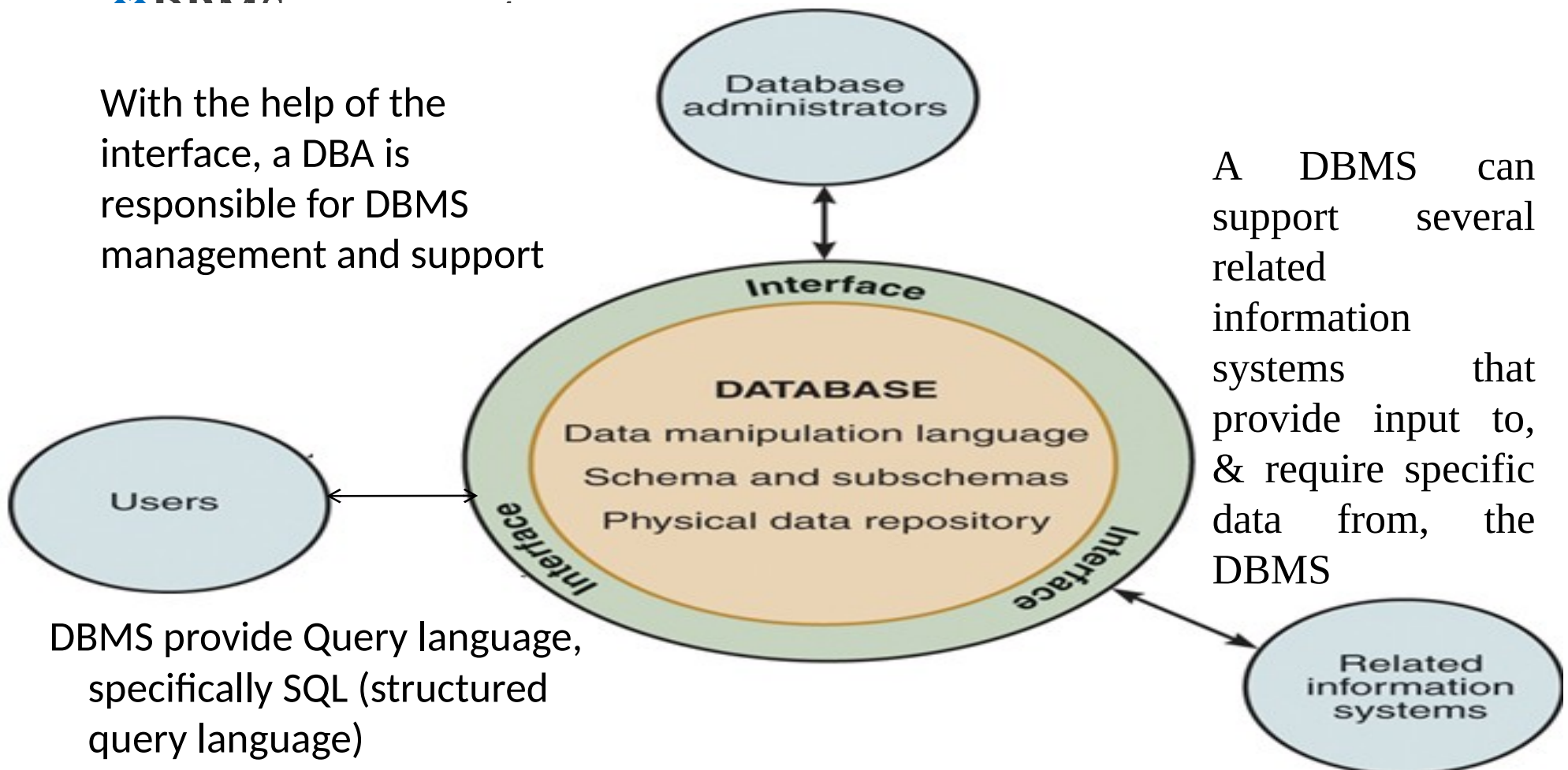
□ What is DBMS?

- ❖ A *Database Management System (DBMS)* is a software package designed to store and manage databases, and controls access to the database.
 - ▶ DBMS is a powerful tool for managing large amounts of data
 - ▶ Provides a systematic method for creating, updating, storing, retrieving data in a database.
- ❖ Popular DBMSs include :
 - ▶ Access, Oracle, MySQL, SQL Server, DB2, Foxpro, etc...
- ❖ A **DBMS** provides an **interface** between a **database** and **users** who need to access the data

Cont...

DBMS

With the help of the interface, a DBA is responsible for DBMS management and support



Cont...

Database Management System (DBMS) Functions

- ▮ The DBMS is an application program that provides an interface for the user to access data as simple as possible. It has several other functions.
 - ❖ **Data storage, retrieval and update:** DBMS allows users to store, retrieve & update data, without knowing the internal structure of the database.
 - ❖ **Creation and maintenance of the data dictionary:** DBMS maintains meta data about the database organization

Cont...

- ❖ **Managing the facilities for sharing the database:** DBMS has to ensure that problems do not arise when two or more users simultaneously access a record & try to update it.
- ❖ **Backup and recovery:** DBMS must provide the ability to recover the database in the event of system failure.
- ❖ **Security:** DBMS must handle users account, & control access to the database based on the privilege users allowed.

Actors on the Database

Database Administrator (DBA)

- ▶ Responsible for managing the database system, authorizing access, coordinating & monitoring uses, acquiring resources (database, DBMS)

Database designers

- ▶ Responsible for designing the database, identifying the data to be stored, choosing the structures to represent and store the data.

Application developers

- ▶ Responsible to write software to allow end users to interface with the database system

End users

- ▶ The persons that use the database for querying, updating, generating reports, etc. to achieve some goal.

Database Development Life-cycle

-  Database planning
-  Data analysis
-  Define Problems and Constraints
-  Database design
-  DBMS selection
-  Implementation
-  Testing
-  Operational maintenance

Database Development Life-cycle(Cont...)

Database Planning

- ❖ A plan is used to provide frameworks that enable the manager to make reasonable estimates of resources, cost and schedule.
- ❖ Planning activities
 - ▶ Effort, cost, resource and duration estimation
 - ▶ Suitable process for the execution
 - ▶ Staff organization
 - ▶ Risk Handling - Identification, analysis and mitigation
 - ▶ Quality assurance plan, configuration management plan, etc...

Data analysis

- ▶ Gathering and developing end user data views.
- ▶ Identify the existing system and desired output
- ▶ Identify the company's business rules and analyze their impacts.

Database Development Life-cycle(Cont...)

□ Define Problems and Constraints

- How does the existing system function?
- What input does the system require?
- What reports does the system generate?
- How is the system output used? By Whom?
- What are the operational relationships among business units?
- What are the limits and constraints imposed on the system?

Database Development Life-cycle(Cont...)

▮ Database Design

- ▮ Gather user and system requirements
- ▮ Create a ***conceptual model*** of the database using the **Entity Relationship model** that is based on the user requirements.
 - ▮ **Conceptual Design:** concise description of
 - ▮ Entities and
 - ▮ Relationship within the target application domain.
- ▮ Convert this **conceptual model (E-R)** into a ***logical database model***
 - ▮ **Logical Design:** used to ensuring via normalization procedures and the definition of integrity rules that the stored data will be non-redundant and properly connected.
 - ▮ Attributes: identifying Primary key and Foreign key
 - ▮ Mapping E-R diagram into relational schema.

Database Development Life-cycle(Cont...)

- Implement the normalized relations as tables in a relational database.
 - **Physical Design:** specifies how database records are stored, accessed and related to ensure adequate performance.
 - Entity names will automatically be Table names
 - Attributes will be column of the respective tables
 - Identify Column data types
- Choose a DBMS :
 - The selection of an appropriate DBMS to support the database application.
 - Also design the user interface and the application programs using the selected DBMS.
 - Example: SQL server or MS-SQL Server.

Database Development Life-cycle(Cont...)

Implementation

- The physical realization of the database and application designs.
 - Use DDL of DBMS to create database schemas and empty database files.
 - Use DML to create any specified user views.

Testing

- The process of executing the application programs with the intent of finding errors.
 - Use carefully planned test strategies and realistic data.
 - Demonstrates that database and application programs *appear* to be working according to requirements.

Database Development Life-cycle(Cont...)

□ Operational Maintenance

- ▶ The process of monitoring and maintaining the system following installation.
- ▶ Monitoring the performance of the system.
- ▶ If performance falls, may require reorganization of the database.
- ▶ Maintaining and upgrading the database application (when required).
- ▶ Incorporating new requirements into the database application.

Database Development Life-cycle(Cont...)

▮ Major reasons for database development runaways

- 👉 Unclear objectives
- 👉 Bad Planning
- 👉 No project management methodology
- 👉 New Technology
- 👉 Insufficient staff

👉 All of these are relate to system management

👉 Effective system management is key to successfully executing a database system

