DATA BASE SYSTEMS

LECTURE 1

PROPOSED BY

DR: ALSHAIMAA MOSTAFA MOHAMMED

CHAPTER 1

INTRODUCTION TO DATABASE SYSTEMS

DATABASE MANAGEMENT SYSTEM (DBMS)

- DBMS is a program or group of programs that work in conjunction with the operating system to create, process, store, retrieve, control and manage the data.
- It acts as an interface between the application program and the data stored in the database.
- It can be defined as a computerized record-keeping system that stores information and allows the users to add, delete, modify, retrieve and update that information.

- The DBMS performs the following five primary functions :

1. **Define, create and organize a database**
2. **Input data**
3. **Process data**
4. **Maintain data integrity and security**
5. **Query database**

1. **Define, create and organize a database :** the DBMS establishes the logical relationships among different data elements in a database and also defines schemas and subschemas using the DDL.
2. **Input data :** it performs the function of entering the data into the database through an input device (like data screen, or voice activated system) with the help of the user.

3. **Process data :** it performs the function of manipulation and processing of the data stored in the database using the DML.
4. **Maintain data integrity and security :** it allows limited access of the database to authorized users to maintain data integrity and security.
5. **Query database :** it provides information to the decision makers that they need to make important decisions. This information is provided by querying the database using SQL.

COMPONENTS OF DBMS

➢ **Data definition language (DDL)**

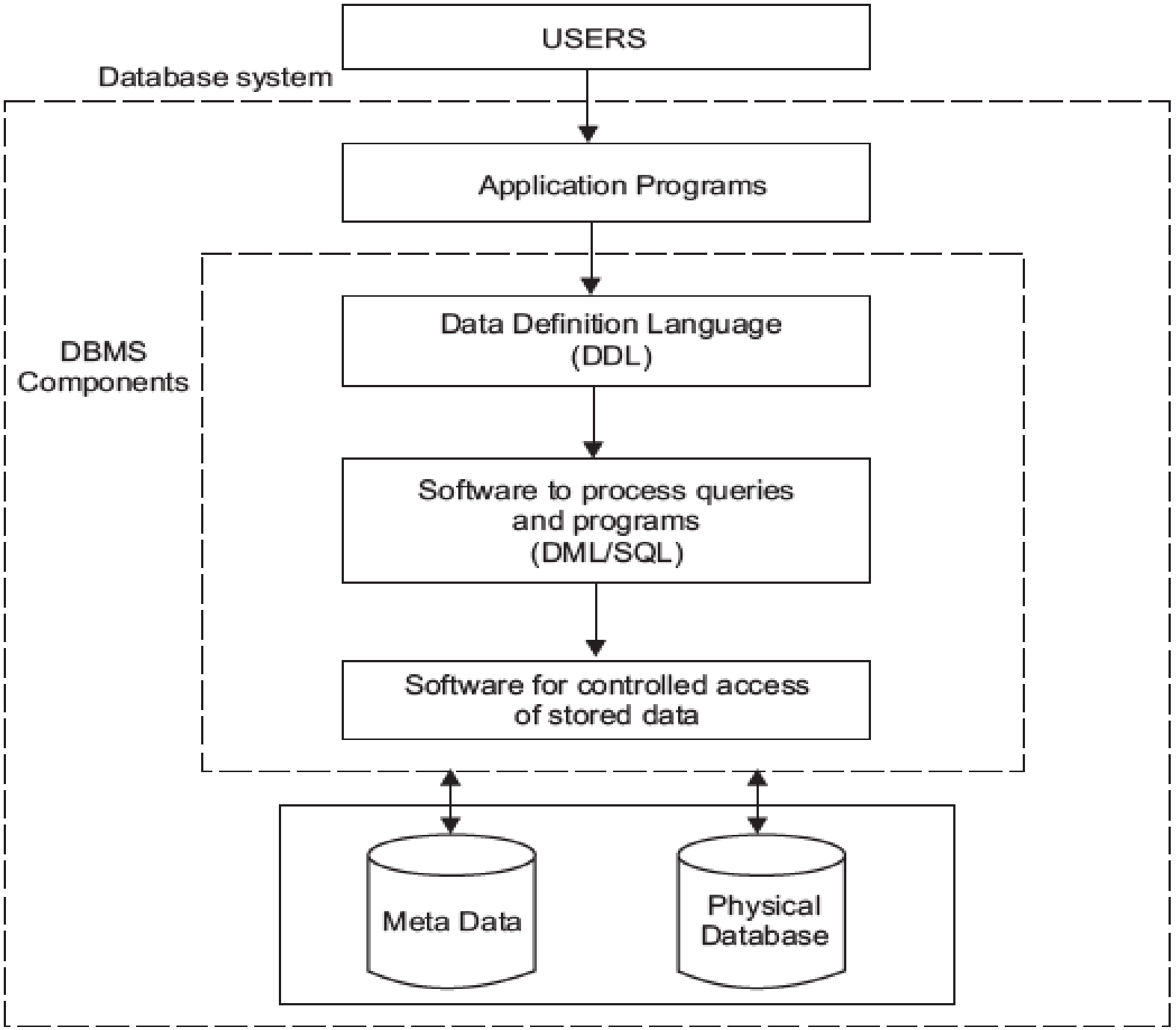
- It allows the users to define the database, specify the data types, data structures and the constraints on the data to be stored in the database.

➢ **Data manipulation language (DML) and query language**

- DML allows users to insert, update, delete and retrieve data from the database. SQL provides general query facility.

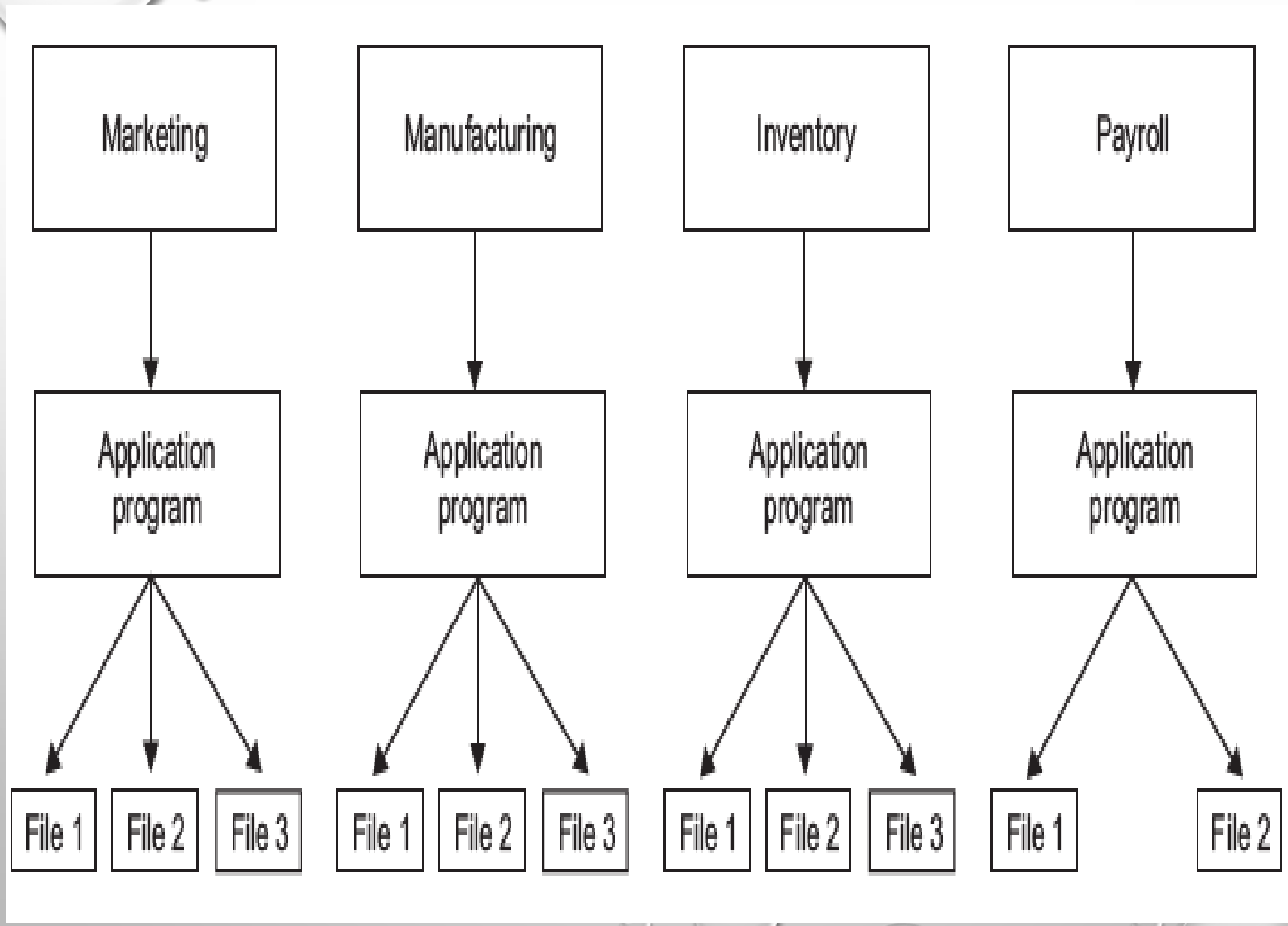
➢ **Software for controlled access of database**

- This software provides the facility of controlled access of the database by the users, concurrency control to allow shared access of the database and a recovery control system to restore the database in case of hardware or software failure.


USERS
Database system
Application Programs
DBMS
Components
Data Definition Language
(DDL)
Software to process queries
and programs
(DML/SQL)
Software for controlled access
of stored data
Meta Data
Physical
Database

TRADITIONAL FILE SYSTEM VERSUS DATABASE SYSTEMS

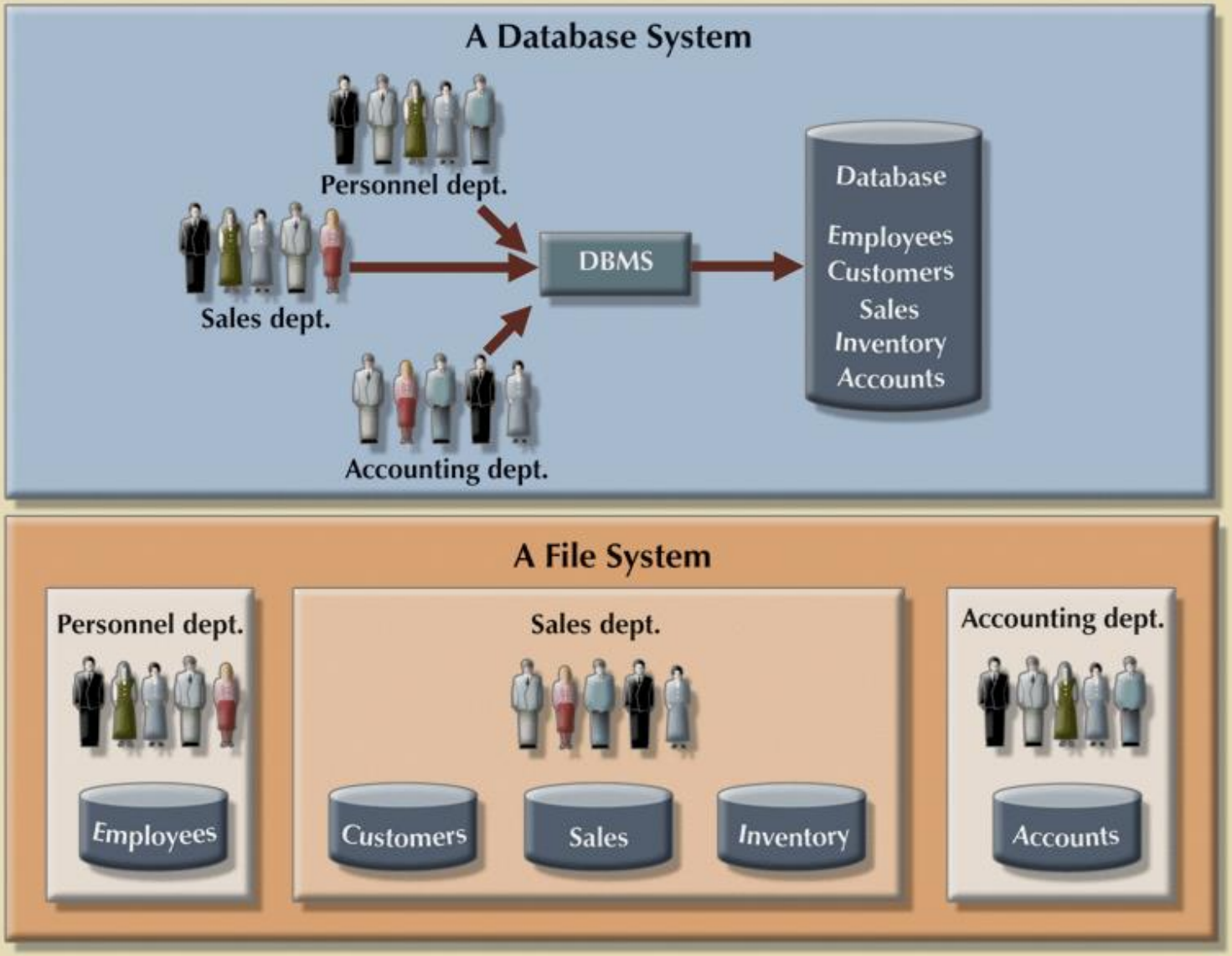
- Each file is independent of other file, and data in different files can be integrated only by writing individual program for each application.
- The data and the application programs that uses the data are so arranged that any change to the data requires modifying all the programs that uses the data. This is because each file is hard-coded with specific information like data type, data size etc.
- Some time it is even not possible to identify all the programs using that data and is identified on a trial-and-error basis.


Marketing
Manufacturing
Inventory
Payroll
Application program
Application program
Application program
Application program
File 1
File 2
File 3
File 1
File 2
File 3
File 1
File 2
File 3
File 1
File 2

DISADVANTAGES OF TRADITIONAL FILE SYSTEM

1. Data redundancy
2. Data inconsistency
3. Lack of data integration
4. Program dependence
5. Data dependence
6. Limited data sharing
7. Poor data control
8. Problem of security

FIGURE 1.6 Contrasting database and file systems



Database Systems or Database System Environment

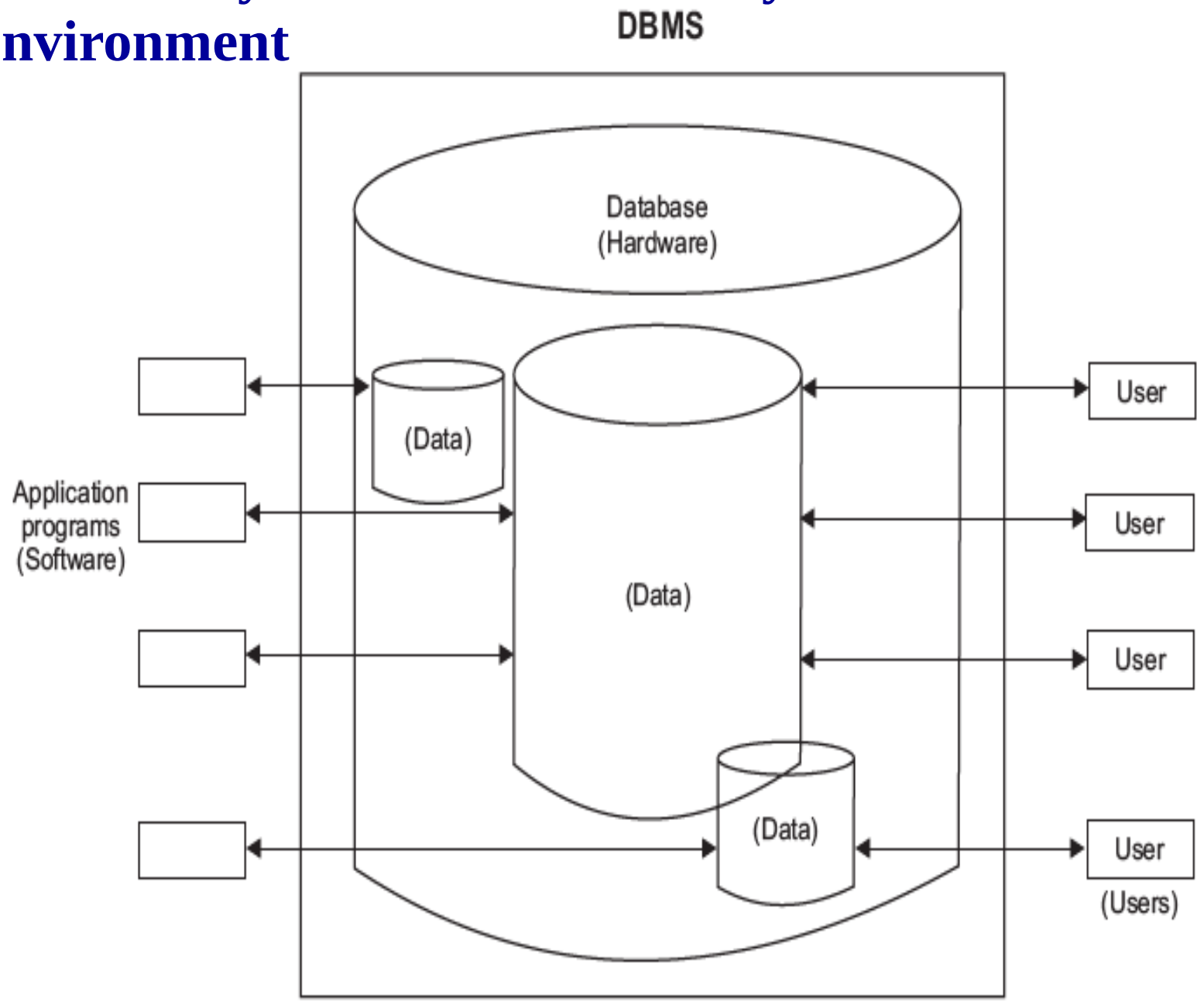
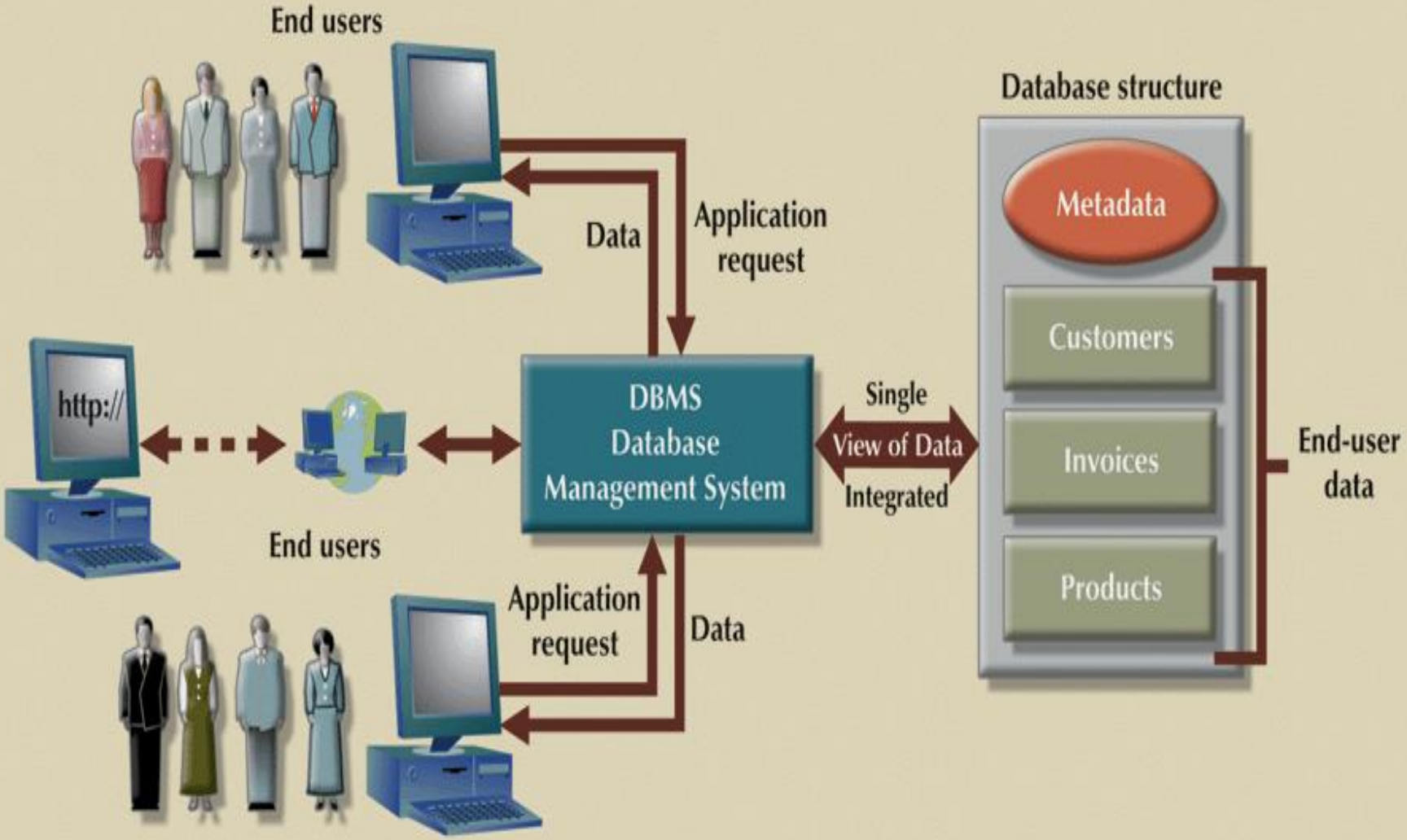


FIGURE 1.2 The DBMS manages the interaction between the end user and the database



DATABASE SYSTEMS OR DATABASE SYSTEM ENVIRONMENT

- **Data** : The whole data in the system is stored in a single database. This data in the database are both shared and integrated. Sharing of data means individual pieces of data in the database is shared among different users and every user can access the same piece of data but may be for different purposes. Integration of data means the database can be function of several distinct files with redundancy controlled among the files.

- **Hardware** : The hardware consists of the secondary storage devices like disks, drums and so on, where the database resides together with other devices. There is two types of hardware. The first one, i.e., processor and main memory that supports in running the DBMS. The second one is the secondary storage devices, i.e., hard disk, magnetic disk etc., that are used to hold the stored data.
- **Software** : A layer or interface of software exists between the physical database and the users. This layer is called the DBMS. All requests from the users to access the database are handled by the DBMS. Thus, the DBMS shields the database users from hardware details. The DBMS provides the other facilities like accessing and updating the data in the files and adding and deleting files itself.

- **Users** : The users are the people interacting with the database system in any way. There are four types of users interacting with the database systems. These are Application Programmers, online users, end users or naive users and finally the Database Administrator (DBA).

END USERS OR NAIVE USERS

- The end users or naive users use the database system through a menu-oriented application program, where the type and range of response is always displayed on the screen.
- The user need not be aware of the presence of the database system and is instructed through each step.

ONLINE USERS

- These type of users communicate with the database directly through an online terminal or indirectly through an application program and user interface.
- They know about the existence of the database system and may have some knowledge about the limited interaction they are permitted.

APPLICATION PROGRAMMERS

- These are the professional programmers or software developers who develop the application programs or user interfaces for the naive and online users.
- These programmers must have the knowledge of programming languages such as assembly, c, c++, java, or sql, etc.,

DATABASE ADMINISTRATOR

- Database administrator (DBA) is a person who have complete control over database of any enterprise.
- Dba is responsible for overall performance of database.
- He is free to take decisions for database and provides technical support.
- He is concerned with the back-end of any project.

MAIN RESPONSIBILITIES OF DBA

- **Deciding the conceptual schema or contents of database :** DBA decides the data fields, tables, queries, data types, attributes, relations, entities or you can say that he is responsible for overall logical design of database.
- **Deciding the internal schema of structure of physical storage :** DBA decides how the data is actually stored at physical storage, how data is represented at physical storage.
- **Deciding users :** DBA gives permission to users to use database. Without having proper permission, no one can access data from database.
- **Deciding user view :** DBA decides different views for different users.

MAIN RESPONSIBILITIES OF DBA

- **Granting of authorities :** DBA decides which user can use which portion of database. DBA gives authorities or rights to data access. User can use only that data on which access right is granted to him.
- **Deciding constraints :** DBA decides various constraints over database for maintaining consistency and validity in database.
- **Security :** security is the major concern in database. DBA takes various steps to make data more secure against various disasters and unauthorized access of data.
- **Monitoring the performance :** DBA is responsible for overall performance of database. DBA regularly monitors the database to maintain its performance and try to improve it.

MAIN RESPONSIBILITIES OF DBA

- **Backup :** DBA takes regular backup of database, so that it can be used during system failure. Backup is also used for checking data for consistency.
- **Removal of dump and maintain free space :** DBA is responsible for removing unnecessary data from storage and maintain enough free space for daily operations. He can also increase storage capacity when necessary.
- **Checks :** DBA also decides various security and validation checks over database to ensure consistency.
- **Liaisioning (communication) with users :** another task of the DBA is to liaisioning with users and ensure the availability of the data they require and write the necessary external schemas.

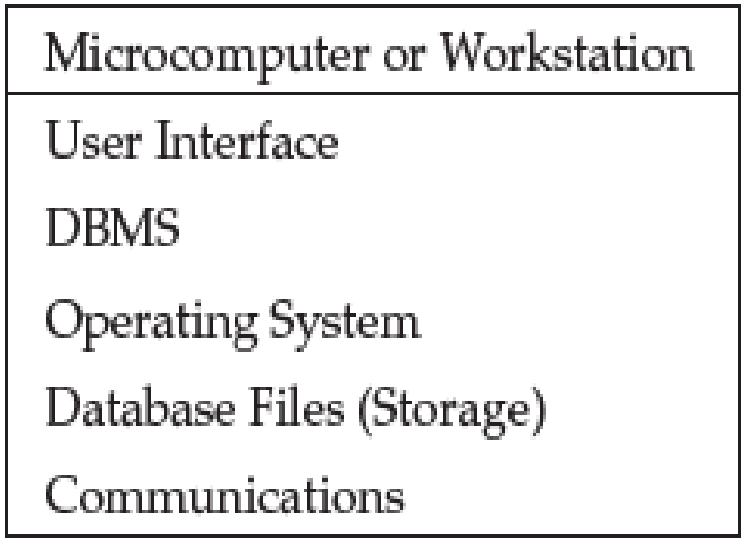
TYPES OF DATABASES

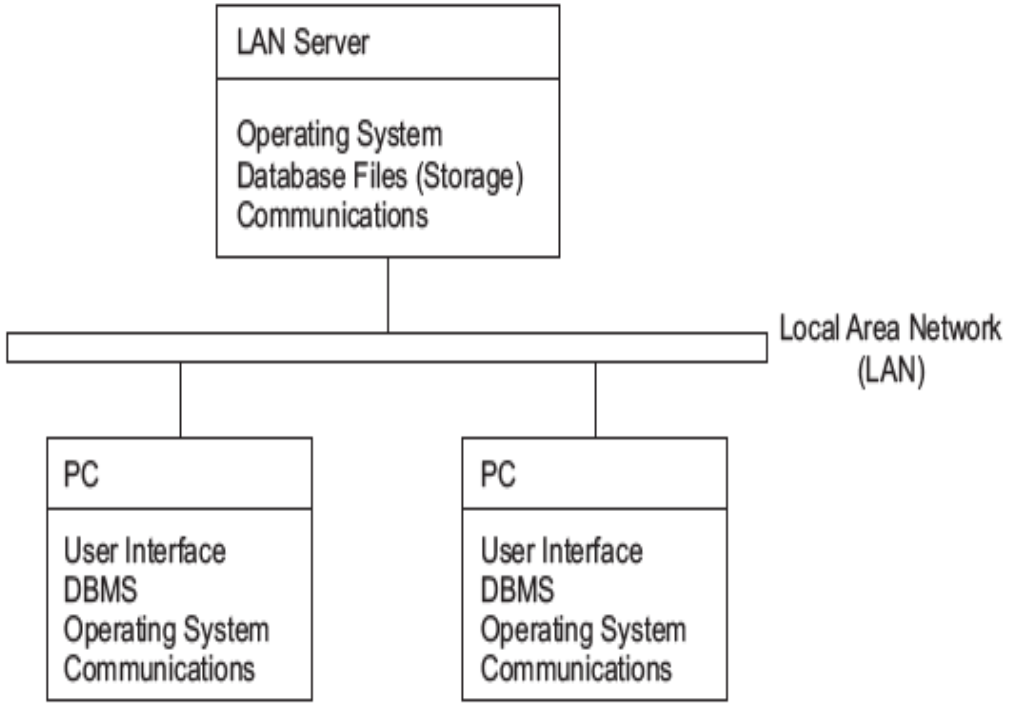
- Databases can be classified according to:
- Number of users
- The type of use
- Database site locations

Number of users

Single-user database systems

Multiuser database systems.





MULTIUSER DATABASE SYSTEM

- **Advantages :**

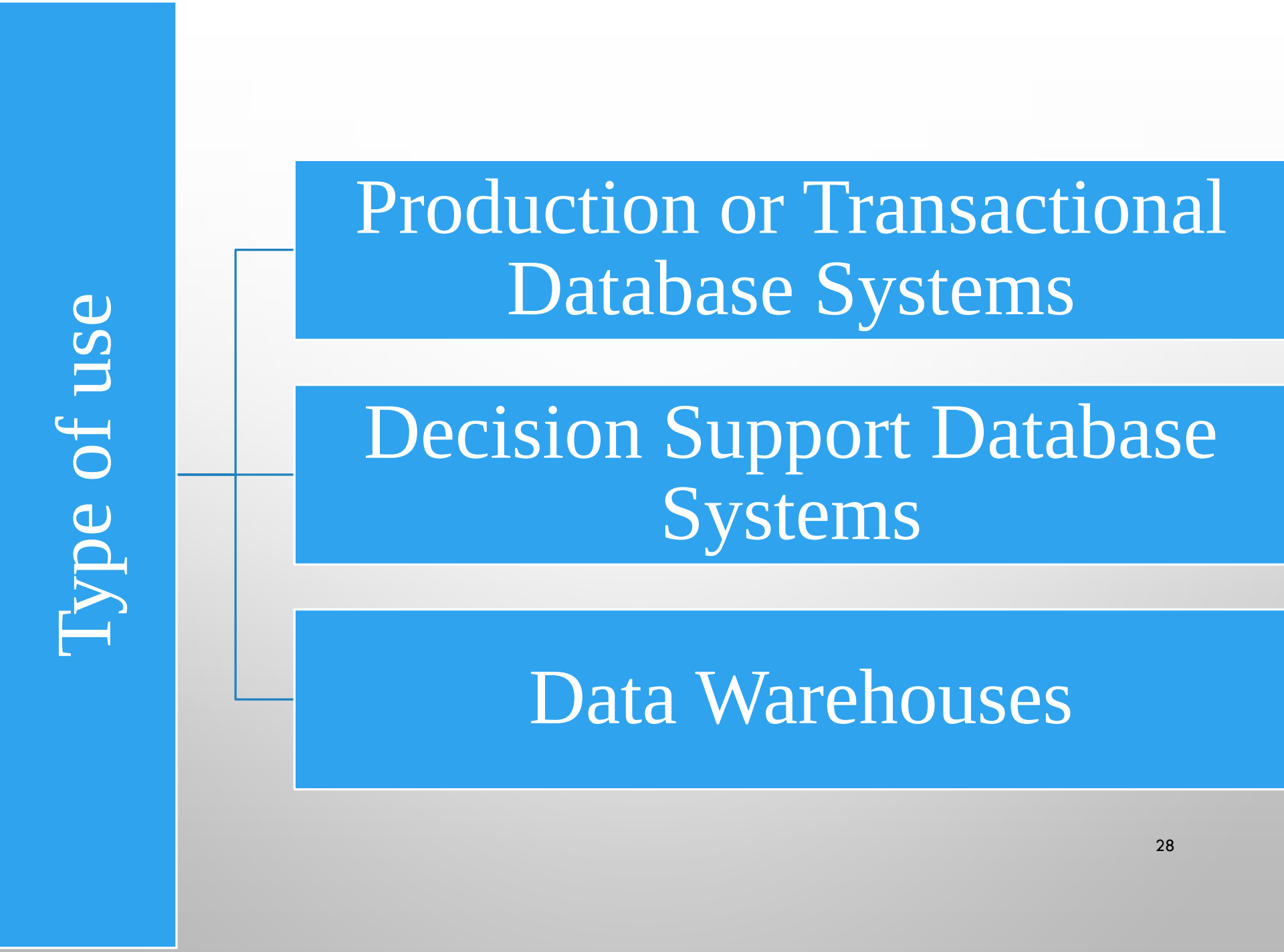
(*i*) ability to share data among various users.

(*Ii*) cost of storage is now divided among various users.

(*Iii*) low cost since most components are now commodity items.

- **Disadvantages :**

The major disadvantage of the multiuser database system is that it has a limited data sharing ability *i.E.,* Only a few users can share the data at most.


Type of use
Production or Transactional Database Systems
Decision Support Database Systems
Data Warehouses

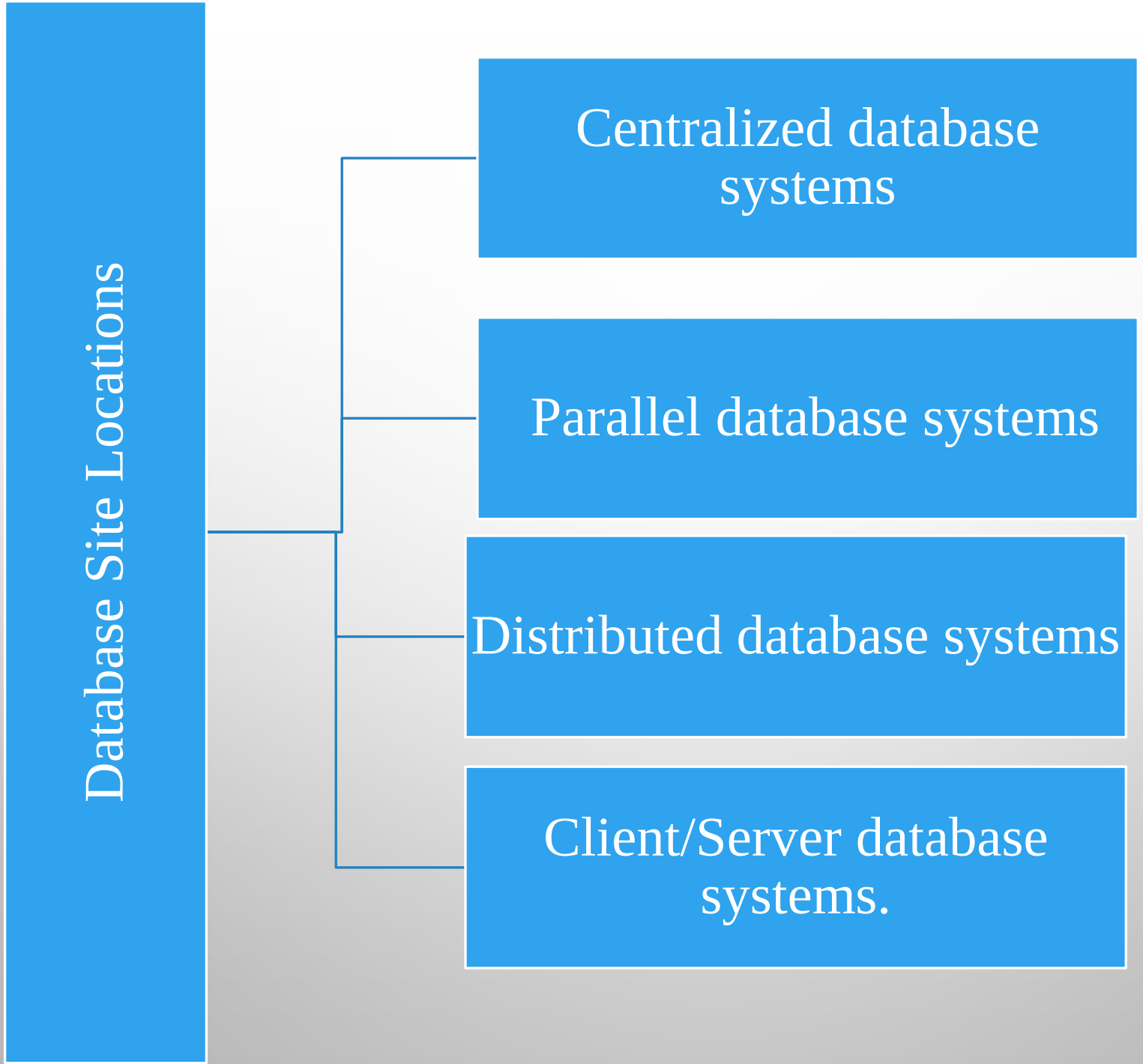

Database Site Locations
Centralized database systems
Parallel database systems
Distributed database systems
Client/Server database systems.

TABLE 1.3

DATABASE CAREER OPPORTUNITIES

JOB TITLE	DESCRIPTION	SAMPLE SKILLS REQUIRED
Database Developer	Create and maintain database-based applications	Programming, database fundamentals, SQL
Database Designer	Design and maintain databases	Systems design, database design, SQL
Database Administrator	Manage and maintain DBMS and databases	Database fundamentals, SQL, vendor courses
Database Analyst	Develop databases for decision support reporting	SQL, query optimization, data warehouses
Database Architect	Design and implementation of database environments (conceptual, logical, and physical)	DBMS fundamentals, data modeling, SQL, hardware knowledge, etc.
Database Consultant	Help companies leverage database technologies to improve business processes and achieve specific goals	Database fundamentals, data modeling, database design, SQL, DBMS, hardware, vendor-specific technologies, etc.
Database Security Officer	Implement security policies for data administration	DBMS fundamentals, database administration, SQL, data security technologies, etc.
Cloud Computing Data Architect	Design and implement the infrastructure for next-generation cloud database systems	Internet technologies, cloud storage technologies, data security, performance tuning, large databases, etc.
Data Scientist	Analyze large amounts of varied data to generate insights, relationships, and predictable behaviors	Data analysis, statistics, advanced mathematics, SQL, programming, data mining, machine learning, data visualization

Summary

- Data are raw facts
- Information is the result of processing data to reveal its meaning
- Accurate, relevant, timely information is the key to good decision making
- Data usually stored in a database
- DBMS implements a database and manages its contents
- Metadata is data about data
- Database design defines the database structure
  - Well-designed database facilitates data management, generates valuable information
  - Poorly-designed database leads to bad decision making, organizational failure
- Databases evolved from manual and computerized file systems
  - In a file system, data stored in independent files
    - Each requires its own management program

Summary

- Some limitations of file system data management:
  - Requires extensive programming
  - System administration complex and difficult
  - Changing existing structures difficult
  - Security features likely inadequate
  - Independent files tend to contain redundant data
    - Structural and data dependency problems
- Database management systems developed to address file system's inherent weaknesses
- DBMS present database to end user as single repository
  - Promotes data sharing
  - Eliminates islands of information
- DBMS enforces data integrity, eliminates redundancy, promotes security

CHAPTER 2

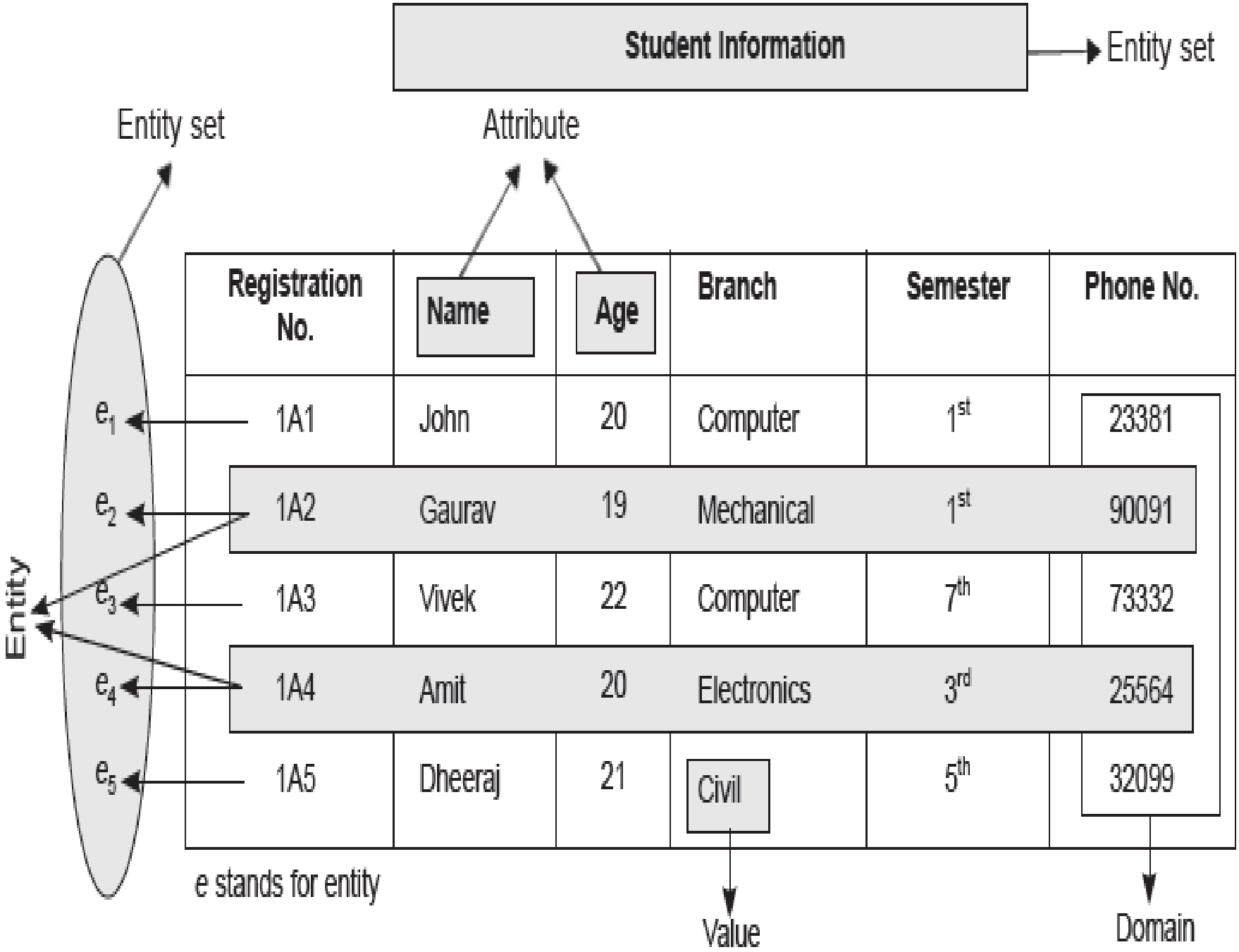
E-R MODELS

OBJECTIVES

- Identify the main characteristics of entity relationship components
- Describe how relationships between entities are defined, refined, and incorporated into the database design process
- See how ERD components affect database design and implementation
- Understand that real-world database design often requires the reconciliation of conflicting goals

The entity relationship (ER) model

- ER model forms the basis of an ER diagram
- ERD represents the conceptual database as viewed by end user
- ERDs depict the ER model’s three main components:
  - Entities
  - Attributes
  - Relationships


Student Information
Entity set
Entity set
Attribute
Registration No.
Name
Age
Branch
Semester
Phone No.
e1
1A1
John
20
Computer
1st
23381
e2
1A2
Gaurav
19
Mechanical
1st
90091
e3
1A3
Vivek
22
Computer
7th
73332
e4
1A4
Amit
20
Electronics
3rd
25564
e5
1A5
Dheeraj
21
Civil
5th
32099
Entity
e stands for entity
Value
Domain

The entity relationship (ER) model

- Enterprise
- Enterprise refers to any kind of organization.
- Ex. Colleges, schools, banks, any company etc.

Entities

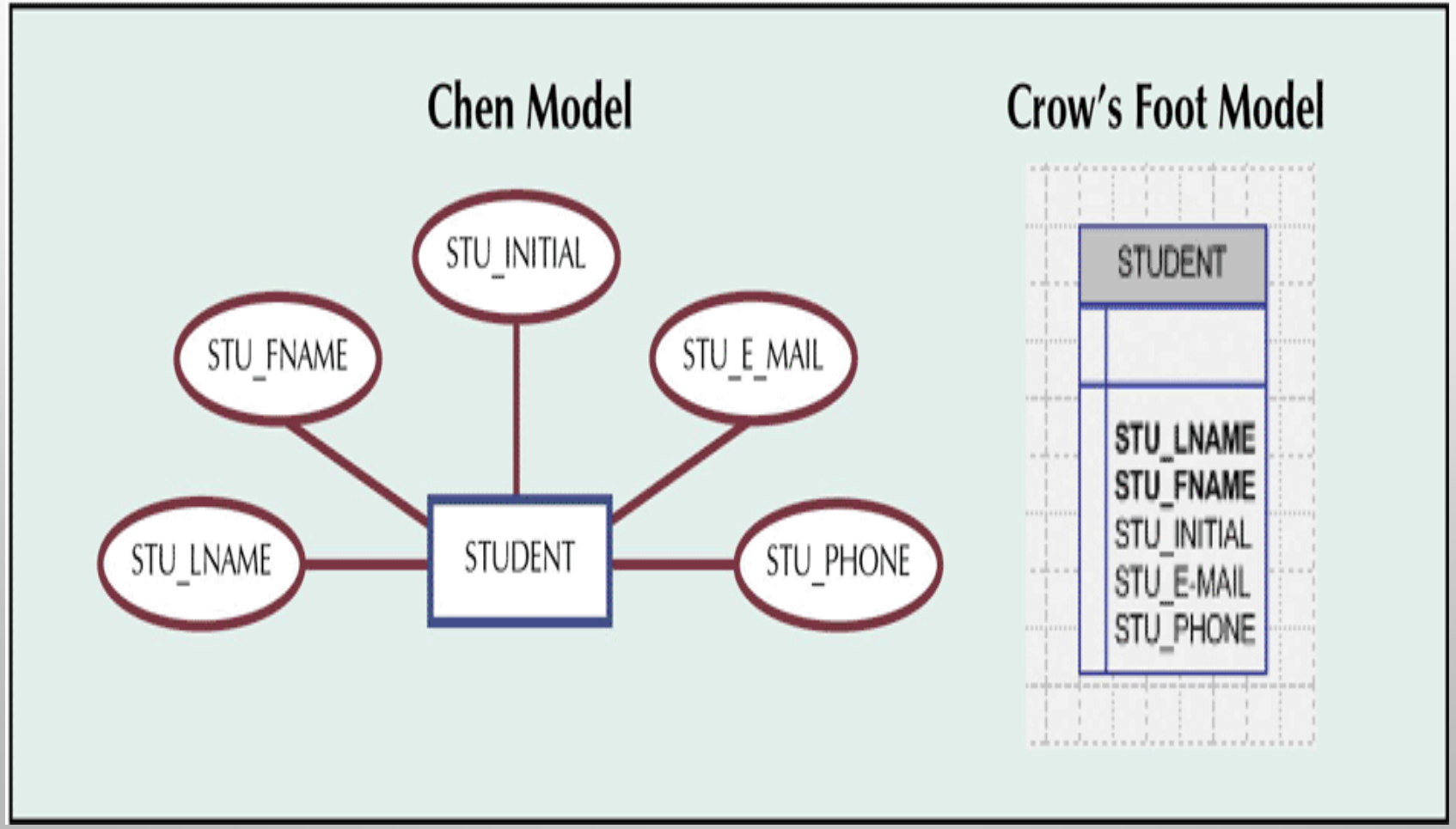
- An entity is an object of interest to the end user.
- Refers to the *entity set* and not to a single entity occurrence
- Corresponds to a table and not to a row in the relational environment
- In both the Chen and Crow's foot models, an entity is represented by a rectangle containing the entity's name
- Entity name, a noun, is usually written in capital letters

Attributes

- Characteristics of entities
- In Chen model, attributes are represented by ovals and are connected to the entity rectangle with a line
- Each oval contains the name of the attribute it represents
- In the Crow’s Foot notation, the attributes are written in the attribute box below the entity rectangle.

The attributes of the STUDENT entity

FIGURE 4.1 THE ATTRIBUTES OF THE STUDENT ENTITY



Domains

- **Attributes have a *domain*:**
  - The attribute’s set of possible values
- **Attributes may share a domain**
  - For example, the address attribute for both customer and agent can have similar type entries

Required and optional attributes

- A required attribute is an attribute that must have a value; in other words, it cannot be left empty
- An optional attribute is an attribute that does not require a value; therefore, it can be left empty.

Primary keys

- Underlined in the ER diagram
- Key attributes are also underlined in a frequently used table structure shorthand
  - Class (**CLASS_CODE**, CRS_CODE, CLASS_SECTION, CLASS_TIME, ROOM_CODE, PROF_NUM)
- Ideally composed of only a single attribute
- Possible to use a *composite key*:
  - Primary key composed of more than one attribute
  - Class (**CRS_CODE, CLASS_SECTION**, CLASS_TIME, ROOM_CODE, PROF_NUM)

The CLASS table (entity) components and contents

FIGURE 4.2 THE CLASS TABLE (ENTITY) COMPONENTS AND CONTENTS

Table name: CLASS Database name: Ch04_TinyCollege

CLASS_CODE	CRS_CODE	CLASS_SECTION	CLASS_TIME	CLASS_ROOM	PROF_NUM
10012	ACCT-211	1	MWF 8:00-8:50 a.m.	BUS311	105
10013	ACCT-211	2	MWF 9:00-9:50 a.m.	BUS200	105
10014	ACCT-211	3	TTh 2:30-3:45 p.m.	BUS252	342
10015	ACCT-212	1	MWF 10:00-10:50 a.m.	BUS311	301
10016	ACCT-212	2	Th 6:00-8:40 p.m.	BUS252	301
10017	CIS-220	1	MWF 9:00-9:50 a.m.	KLR209	228
10018	CIS-220	2	MWF 9:00-9:50 a.m.	KLR211	114
10019	CIS-220	3	MWF 10:00-10:50 a.m.	KLR209	228
10020	CIS-420	1	W 6:00-8:40 p.m.	KLR209	162
10021	QM-261	1	MWF 8:00-8:50 a.m.	KLR200	114
10022	QM-261	2	TTh 1:00-2:15 p.m.	KLR200	114
10023	QM-362	1	MWF 11:00-11:50 a.m.	KLR200	162
10024	QM-362	2	TTh 2:30-3:45 p.m.	KLR200	162

KEYS

- **keys** are important because they are used to ensure that each row in a table is uniquely identifiable
- They are also used to establish relationships among tables and to ensure the integrity of the data.
- A **key** consists of one or more attributes that determine other attributes.
- Candidate key, foreign key, primary key (PK), secondary key, and key.
- **primary key** (PK) In the relational model, an identifier composed of one or more attributes that uniquely identifies a row.
- Key's role is based on **determination**
- Determination is the state in which knowing the value of one attribute makes it possible to determine the value of another.
- The idea of determination is not unique to the database environment.

Types of keys

- A composite key is a key that is composed of more than one attribute.
- An attribute that is a part of a key is called a key attribute.
- superkey is a key that can uniquely identify any row in the table. In other words,

a superkey functionally determines every attribute in the row.

- A candidate key is a minimal superkey—that is, a superkey without any unnecessary attributes. A candidate key is based on a full functional dependency.
- Entity integrity is the condition in which each row (entity instance) in the table has Its own unique identity.
- To ensure entity integrity, the primary key has two requirements:
  (1) all of the values in the primary key must be unique and
  (2) no key attribute in the primary key can contain a null
- Null values are problematic in the relational model.
- A null is the absence of any data value, and it is never allowed in any part of the primary key.

- A foreign key (FK) is the primary key of one table that has been placed into another table to create a common attribute.
- Foreign keys are used to ensure referential integrity, the condition in which every reference to an entity instance by another entity instance is valid.
- In other words, every foreign key entry must either be null or a valid value in the primary key of the related table.
- Secondary key is defined as a key that is used strictly for data retrieval purposes.