Unit ⇒ 1

Datbase Management System (DBMS)

1.1 Introduction: Data, Information, Database and DBMS

Data: Data is the raw fact which is composed of alphabets, digits, special, symbols, etc. For example, Ram, 26, Mahendranagar, etc.

Information: The processed data that gives meaningful result is called information. It is an organised collection of related data which gives a complete sense. For example, Ram is 26 years old who lives in Mahendranagar.

Database: It is a collection of interrelated data of objects or entities in a structured form / tabular form. For example, customer record, books library, etc.

INSPECTED

CamScanner

DBMS: It is a software that defines, manipulates and manages the database. It allows to access the fields, update the records and retrieve the data as requested. Some examples are Ms-Access, Oracle, My SQL, etc.

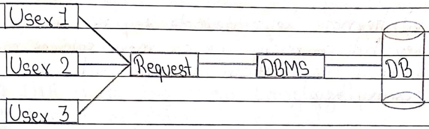


Fig:- Logical view of database system

1.2. Some basic terms in DBMS

i) Field: A field is the property or attributes of a table. For example, name, phone number, etc.

ii) Record: Record is a collection of interrelated fields. For example, 1, Ram, etc.

iii) Table: Table is the arrangement of rows and columns. In table, a row defines a record and column defines field and vice versa.

iv) Objects: They are used in database to store the data. It holds the data for manipulation and display it in the proper format.

CS CamScanner

v) Primary key: Primary key is a unique key/value in database which is used to show the relationship between different tables or objects and also helps in reducing data duplicacy. For example, registration number, citizenship number, etc.

vi) Alternate key: Alternate key is not a primary key but combining one or more columns can make it unique. It can also be used instead of primary key.

vii) Candidate key: Sometimes there is more than one attribute processing the unique identification property. In this case all attribute combination provide a unique record which is called candidate key.

1.3 Advantages of DBMS

Some of the advantages of DBMS are mentioned below:

① Reduce data redundancy/duplicacy.
② Fast and easy file sharing.
③ Data integration.
④ Backup and recovery.
⑤ High data security (Authenticated data)
⑥ Multiple user interface.

CS CamScanner

Disadvantages of DBMS

Some disadvantages of DBMS are mentioned below:

① Setup cost is expensive.
② Changing technology.
③ Needs techical training.
④ Hacking and data leakage.

1.4 DDL (Data Definition Language) and DML (Data Manipulation Language)

DDL is used to specify the content and structure of the table. It defines the physical characteristics of record. It uses command such as CREATE, DROP, RENAME, etc.

Syntax:

```
CREATE table table_name(field1 data_type, field
2 data_type, ...);
```

For example,

```
CREATE table Stu (S_id int (4), S_name char(64)
...);
```

DML is the manipulation of records such as retrieval, sorting, display, deletion, etc. It helps users to use query and display record/report accordingly. It provides techniques to manipulate the database. It uses symbols/commands such as INSERT, DELETE, SELECT, UPDATE, etc.

CS CamScanner

Syntax:

```
INSERT into table name
VALUES (List of values);
```

For example,

```
INSERT into Stu
VALUES (14, "Ram");
```

Differences between DDL and DML

DDL	DML
It is a type of SQL command that helps to define database schemes.	It is a type of SQL command that helps to retrieve and manage data in relational databases.

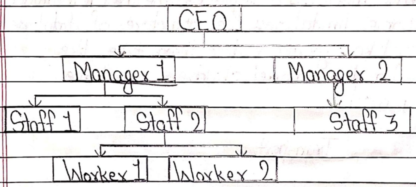
It stands for Data Definition Language	It stands for Data Manipulation Language.
CREATE, DROP, ALTER are some DDL commands	INSERT, UPDATE, DELETE are some DML commands.
Commands affect the entire database or the table.	Commands affect one or more records in a table.
SQL statements cannot be rolled back.	SQL statements can be rolled back.

CS CamScanner

1.5 Database model

There are different forms of database management system, they are:

① ~~Hie~~ Hierarchical model:- It is one of the oldest type of database model in which all records are arranged in a tree like structure. In this model, the records are in a parent-child relationship basis in which children can have only one parent but parents can have many children. And also the top level record gets higher priority than the child at bottom level. For example,



Advantages of hierarchical model

The advantages of hierarchical model are mentioned below:

i) Easiest model to implement.
ii) It supports one-to-one and one-to-many relationship.
iii) Highly secure model.
iv) Searching is easy if the parent is known.

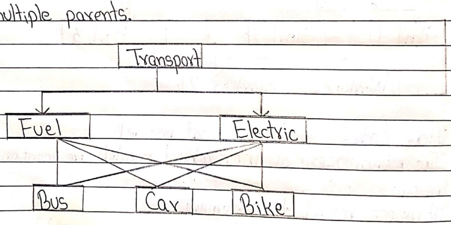
CS CamScanner

Disadvantages of hierarchical model

The disadvantages of hierarchical model are as follows:

i) It is old fashioned and outdated.
ii) It does not support many-to-one relationship.
iii) It increases data duplicacy.
iv) Dependency on parent node.
v) If parent node is disturbed, then the whole hierarchical model may be disturbed.

② Network model:- It replaces hierarchical model due to some limitations. The structure of database is more like graph rather than tree like structure. In this model, children can have multiple parents.



Advantages of network model

The advantages of network model are mentioned below:

i) It supports many-to-one relationship.
ii) Searching is faster.
iii) Simple and easy to design.

CS CamScanner

iv) It reduces data duplicacy.
v) Good for dynamic database.

Disadvantages of network model

The disadvantages of network model are mentioned below:

i) It is less secure.
ii) Difficult for complex relationship.

③ Relational database model:- The database model which stores and displays data in tabular form (i.e. rows and columns) is known as relational database model. For example, MySQL, Oracle, Ms-Access, etc.

RN	Name
1.	Ram
2.	Shyam

RN	Name	Ph.	Dist
1.	Ram	98	Kanch.
2.	Shyam	97	Kail.

SN	Country	Dist.
1	Nepal	Kanch.
2.	Nepal	Kail.

RN	Name	Phone	Dist.	Country
1.	Ram	98	Kanch.	Nepal
2.	Shyam	97	Kail.	Nepal

Advantages of relational database model

The advantages of relational database model are as follows:

① Database processing is faster.
② Low chances of data duplicacy.
③ Data integrity can be easily implemented.

CamScanner

④ Normalization can be done with it.

Disadvantages of relational database model

The disadvantages of relational database model are as follows:

i) More complex than other models.
ii) Difficult to understand.
iii) Sometimes too many roles may be created because of complex relationship.

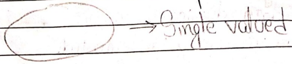
④ ER model:- It shows the relationship between various real life entities and their attributes.

Components of ER diagram

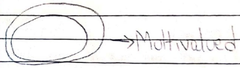
The components of ER diagram are explained below:

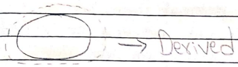
i) Entity:- Entity represents real life objects. For example, student, library, hospital, etc. It is represented by rectangular shape. Its symbol is

ii) Attributes:- Attributes possess the characteristics of entity. It may be single valued, multi-valued or derived. It is represented by oval shape.

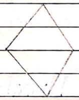


CamScanner





iii) Relation:- Relation shows the logical coordination between two or more than two entities. It is represented by diamond shape.
i.e.



iv) Lines:- Lines are also called connectors that connect entities, attributes and relations. It is represented by single line.
i.e.

For example:

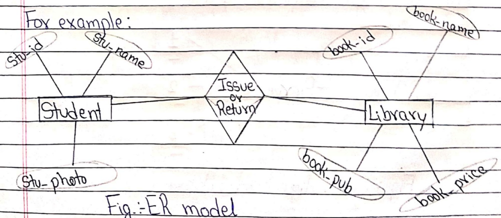


Fig:- ER model

CS CamScanner

Advantages of ER model

The advantages of ER model are mentioned below:

i) Easy to represent and understand the flow of data as it is in pictorial form.
ii) Helps in creating a link between several entities.
iii) Easy to implement ER-diagram into DBMS.

Disadvantages of ER-model

The disadvantages of ER-model are mentioned below:

i) It may become complex with more no. of relations.
ii) Users need to remember symbols.

1.6 Centralized vs Distributed database

① Centralized database:- It is a simple type of database which works on client server basis. In this type, the database is centralized and run on a single computer system.

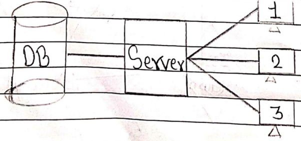


Fig.:- Centralized database

CamScanner

Advantages of centralized database

The advantages of centralized database are as follows:

i) It is easy to implement.
ii) Cost friendly.
iii) Suitable for smaller organization.
iv) High security.
v) Maintainance of database is easier.

Disadvantages of centralized database

The disadvantages of centralized database are as follows:

i) It is dependent on central server. If the server goes down, the whole system will be disturbed.
ii) Not suitable for large organization.
iii) Not suitable for huge data traffic.

② Distributed database:- It is a complex type of database system which is a collection of multiple logically interrelated databases which may be in different geographical locations. In other words, it is a collection of different centralized databases in different locations.

Advantages of distributed database

The advantages of distributed database are as follows:

i) It can handle huge data traffic at the same time.
ii) Backup and recovery is easier.
iii) It is suitable for large organization.

CamScanner

Disadvantages of distributed database

The disadvantages of distributed database are mentioned below:

i) It is higly expensive.
ii) It is highly complex.
iii) Security issues.
iv) Sometimes there may be collision of data.
v) Maintanance is difficult.

Differences between centralized database and distributed database

Centralized database	Distributed database
It depends upon central server.	It has several servers.
It can't handle huge traffic at the same time.	It can handle huge traffic at the same time.
Maintainance is easier.	Maintainance is difficult.
More secure.	Less secure.
Not suitable for large organization.	Suitable for large organization.

CS CamScanner

1.7 Concept of Normalization

Normalization is a process in which complex database table is broken down into separate tables. This makes data model more flexible and easier to maintain. Hence it reduces complexity of database.

Types of normalization

The types of normalization are explained below:

i) 1NF : First normal form (1NF) sets the very basic rule for an organized database.
→ The datafield must be atomic (single record).
→ It eliminates duplicate rows and columns of the same table.
→ It minimizes the data redunancy in the database table.

For example,

Random data

SN	Name	Phone No
1.	Ram Bhatt	9848
		9742
2.	Sita Pant	9785

First normal form

SN	first_name	last_name	Phone No
1	Ram	Bhatt	9848
1	Ram	Bhatt	9742
2	Sita	Pant	9785

CamScanner

ii) 2NF :- Second normal form (2NF) further addresses the concept of removing duplicate data. To be in 2NF, the table must be in 2NF. It defines data dependency. Here, non-key attributes are functionally depend on key attributes (or primary attributes).

For example,

SN	Name	Phone No
1	Ram Bhatt	9848
		9742
2	Sita Pant	9785

SN	first_name	last_name
1	Ram	Bhatt
2	Sita	Pant

SN	Phone No.
1	9848
1	9742
2	9785

2NF

Primary key

iii) 3NF :- To be in third normal form (3NF), the data must be in 2NF and it must remove transitive dependency. Here, transitive dependency means, if A→B and B→C, then A→C.

For example,

SN	Name	City	District	Country
1.	Ram	MNR	Kanchanpur	Nepal
2.	Shyam	Banbasa	Uttarakhand	India
3.	Hari	Kapan	Kathmandu	Nepal

CamScanner

SN	Name	City
1	Ram	MNR
2	Shyam	Banbasa
3	Hari	Kapan

City	District
MNR	Kanchanpur
Bahbasa	Uttarakhand
Kapan	KTM

District	Country
Kanchanpur	Nepal
Uttarakhand	India
KTM	Nepal

1.8 Data Security

Data security refers to the safety of hardware and software data required for DBMS.

Hardware safety/security

i) Regular maintenance of the hardware components.
ii) Temperature control
iii) Free from dust
iv) Electric stabilization

Software security

i) Use genuine software
ii) Use strong anti-virus and update it regularly
iii) Use of strong firewall
iv) Use of proper authentication techniques pattern, PIN, biometric, OTP, etc.
v) Use VPN if necessary.

CamScanner