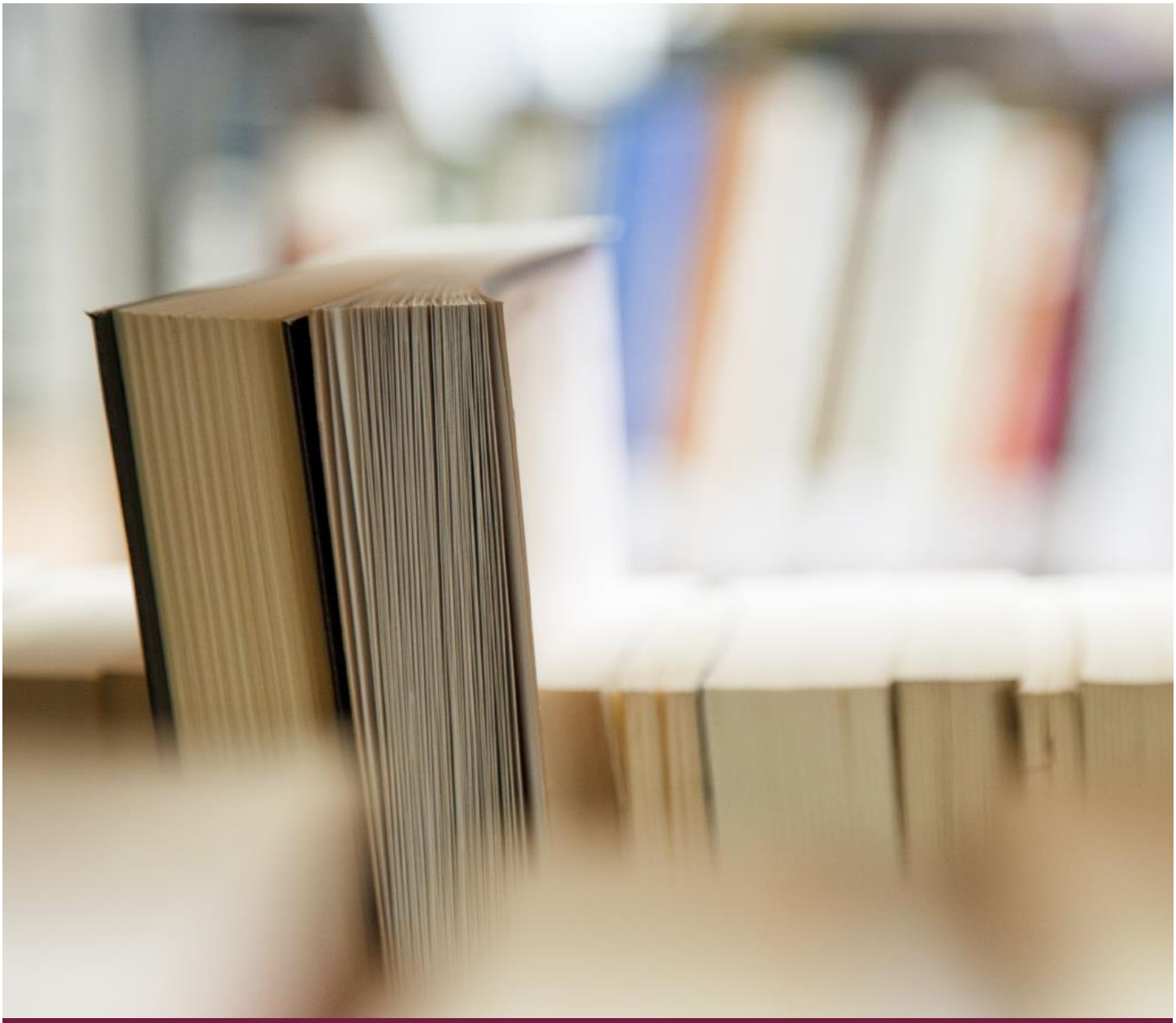


UNIT-01

DBMS

Prof. Akshara V. Chavan SYKKSU

DATA

- Data are facts that can be collected through observation & measurement
- For example: amount, roll no, name, telephone no etc.
- The data are organized in the form of characters, fields, record, files and databases. There are 2 types of data:
  - It is the collection of information needed by the organization.
  - Metadata - is information about the data i.e. data about data**.**

DATABASE

- A database is a collection of data elements (facts) stored in a computer in a systematic way.

DATABASE MANAGEMENT SYSTEM(DBMS)

- A Database Management System (DBMS) is a software package to facilitate the creation and maintenance of a computerized database.
  - Functionalities related to DBMS are:
  - Creation of database
  - Insertion of data into DB
  - Updating the data in DB
  - Effective retrieval of data from DB
  - Securing of data stored in DB
  - Deletion of data from DB
- The Database Management System is largely called as DBMS. This DBMS basically contains three important terms
- Database: Storage of data in large quantity
- Management: Database handling activities
  - Creation of database
  - Insertion of data into DB
  - Updating the data in DB

- Effective retrieval of data from DB
- Securing of data stored in DB
- Deletion of data from DB

- System: A complex program or software

> Complex software that performs management activities like "creation, addition of data, modifying, removing data and accessing data" on a database is called as Database Management System

INTRODUCTION TO DATABASE

- Database is a storage that holds large amount of data.
- The stored data is used to produce information as and when required.
- Generally speaking, the database is nothing but a storage place of data which is of interest to the enterprise.
- The database is physically present on the secondary storage or disk. So, the data stored in the database remains permanent so that it can be used later. The quantity of data stored in database depends on the size of the enterprise.

THE FUNCTIONALITIES RELATED TO THE DATABASE ARE

1. Creation of the database

- Creation of the database starts from the planning.
- When the idea of preparing to store the data arrives, then a model is prepared.
- On the basis of this model preparation of database is done. Preparing this physically is called as creation of data- base.
- The database created contains only the structure or framework. The created database is ready to hold the data.

2. Insertion of data into the database

- Once the database is created the next function is to put data into it. Entering data into the database is done through the insertion operation. So, putting data into the database is called insertion of data.

3. Updating data in the database

- The data stored once may be the basic facts of the activities of the enterprise. Some data need to be derived from these existing data. Some may require change as per the operations. All these activities come under updating functionality of the database.

4. Effective retrieval of data from the database

- Once the data is finalized in the database then the functionality which arrives is its retrieval.
- The data stored in the database is retrieved for the purpose of producing the information.
- The retrieval of data must be such that the response time must be as low as possible.

Such retrieval is called as effective data retrieval functionality of the database

5. **Securing of data stored in the database**
   - The data stored in the database is used by many people of the enterprise. So, there is every chance losing authenticity of the data.
   - The permitted users of the database must access the data. It is necessary that proper security mechanisms must be used to ensure it. This facility of providing security is called securing the data functionality of the database.
6. **Deletion of data from the database**
   - When the data is of no use to the enterprise, that is it has become obsolete, it must be removed from the database. Removing the old data from the database is called as deletion of data functionality of the database.

PURPOSE OF DATABASE SYSTEM

1. DATA REDUNDANCY

- Data redundancy refers to the duplication of data.
- Let's say we are managing the data of a college where a student is enrolled for two courses, the same student details in such case will be stored twice, which will take more storage than needed.
- Data redundancy often leads to higher storage costs and poor access time.

2. DATA INCONSISTENCY

- Data redundancy leads to data inconsistency
- In file processing system, various copies of same data may contain different values. Data is not consistent in this system, it means if a data item needs to be changed then all the files containing that data need to be modified. It may create a risk of out dated values of data.

3. ATOMICITY ISSUES

- Atomicity is required to save the data values, it means that information is completely entered or canceled at all. Any system may fail at any time and at that time it is desired that data should be in a consistent state.
- Atomicity of a transaction refers to "All or nothing", which means either all the operations in a transaction executes or none.

4. POOR DATA SECURITY

- Poor data security is the most threatening problem in File Processing System. There is very less security in File Processing System as anyone can easily modify and change the data stored in the files.
- All the users must have some restriction of accessing data up to a level.

5. WASTAGE OF LABOR AND SPACE

- Labor is very costly in this era and no organization can afford wastage of their precious labor.
- File Processing System needs lots of copied data in different files that cause wastage of labor. Also maintaining same data again and again leads to wastage of space too.

6. UNSTANDARDIZED DATA

- A collection of data is integrated if it meets certain consistency constraints. A programmer always puts these constraints in the programs by adding some codes.
- In File Processing System, poor data integrity often arises and it becomes very difficult to add new constraints at that time.

ADVANTAGES OF DATABASE MANAGEMENT SYSTEM

1. CENTRALIZED MANAGEMENT OF DATA

- The database management system (DBMS) helps in placing every data of the enterprise at only one place centrally in the form of a database

2. SHARING OF DATA

- When the data is stored centrally stored, it can be shared among many people of the enterprise or organization or institution depending on their requirements.
- The Director, the Managers, the Supervisors, the Workers, the Customers and the Investors can access the database for their use but in restricted or controlled manner

3. INDEPENDENCE OF DATA

- The data stored in the database of DBMS is independent at various levels of access.
- The application programmers or users view the data in different way. It is external to them.
- Their view of data is independent of how logically the database is prepared by the database designer. It is logical data independence.
- For the database designer as well as external users again, it is independent the way how it is physically stored on the secondary storage or physical device used. It is physical data independence.

4. NON-REDUNDANCY OF DATA

- Repetition of the same data at more than one place is called as data redundancy.
- For the sake to avoid some loss of information, the data can be duplicated by the database designer as per the requirement.
- The process of removing the redundancy of data is called as non-redundancy of data. It helps in saving the physical storage space, processing time as well as communication cost.

5. EASE OF DATA ACCESS

- As the data is stored at one place it can be accessed easily by all the users of database.
- This access can be granted and controlled by the administrator depending on the user and his level of access.
- Efficient techniques can be imposed for effective retrieval of data very easily. This ease of access helps in providing a good response time as well as increased through put.

6. DATA CONSISTENCY AND INTEGRITY

- Consistency in database systems refers to the requirement that and given database transaction must change affected data only in allowed ways.

7. IMPOSING PROPER SECURITY

- The data in the database can be clearly secured in better way at various levels.
- Both internal as well as external security on database can be easily imposed. The DBMS provides security at database level.

8. EASE OF RECOVERY

- DBMS provides techniques to recover the data stored in database when system crashes because of some failures. It is very easy for the administrator to recover the data whenever it is required.

9. SIMULTANEOUS ACCESS

- As the data is shared by many people, it can be even accessed simultaneously. Such simultaneous access is called as concurrent access.
- By means of some controlling measures consistency of the data can be maintained. Such simultaneous access increases the speed of transaction processing.

DISADVANTAGES OF DBMS

1. IMPLEMENTATION COST

- The cost incurred in the development and purchase of software is much higher in comparison to convention file-handling.
- It is necessary to upgrade the hardware accordingly which is again a costly affair. The cost is also incurred in providing workspace for the execution of complex programs along with storage cost.

2. PROCESSING OVERHEAD TO IMPLEMENT SECURITY AND INTEGRITY

- The security and the integrity are the major requirements of DBMS that prove the worthiness of its implementation.
- Implementation of proper security and integrity involves more processing overhead on the DBMS.

3. DEGRADATION OF RESPONSE TIME AND THROUGHPUT

- The database is being shared by many users which is the major advantage of DBMS.
- The time required to get the result or data from the DBMS is called as response time. When more users attempt for the same database for the data items, the response time is quite large. It is fast in one-to-one access.
- When the response time is large the throughput is reduced. Number of users getting response in a given unit of time from the DBMS is called as throughput.

4. CHANGE OVER COST

- The users of conventional file-handling take time in responding to the new system. When DBMS is new to the enterprise, the people of the enterprise spend much time in adopting to it.

5. BACKUP REQUIREMENT

- To reduce the duplication of data the complete data of an enterprise is centrally placed as at one place of course using a single database.
- This forces the organization to prepare a backup of the database keeping the failure of the system in mind.
- In case of failure the database can be recovered from the backup. In DBMS the recovery and backup are more complex.

6. DOWN TIMES AND FAILURES

- The data in the enterprise is centrally organized and placed at one location as database. So, the enterprise must bear the cost at down times of DBMS and during the failure of DBMS till the recovery. Naturally even the computing system also tends to some failure.

APPLICATION OF DBMS

There are different fields where a database management system is utilized. Following are a few applications which utilize the information base administration framework –

1. RAILWAY RESERVATION SYSTEM

- In the rail route reservation framework, the information base is needed to store the record or information of ticket appointments, status about train's appearance, and flight.
- Additionally, if trains get late, individuals become acquainted with it through the information base update.

2. LIBRARY MANAGEMENT SYSTEM

- There are lots of books in the library so; it is difficult to store the record of the relative multitude of books in a register or duplicate.

- Along these lines, the data set administration framework (DBMS) is utilized to keep up all the data identified with the name of the book, issue date, accessibility of the book, and its writer.

3. BANKING

- Database the executive's framework is utilized to store the exchange data of the client in the information base.

4. EDUCATION SECTOR

- Presently, assessments are led online by numerous schools and colleges. They deal with all assessment information through the data set administration framework (DBMS).
- In spite of that understudy's enlistments subtleties, grades, courses, expense, participation, results, and so forth all the data is put away in the information base.

5. CREDIT CARD EXCHANGES

- The database Management framework is utilized for buying on charge cards and age of month to month proclamations.

6. SOCIAL MEDIA SITES

- We all utilization of online media sites to associate with companions and to impart our perspectives to the world.
- Every day, many people group pursue these online media accounts like Pinterest, Facebook, Twitter, and Google in addition to.
- By the utilization of the data set administration framework, all the data of clients are put away in the information base and, we become ready to interface with others.

7. BROADCAST COMMUNICATIONS

- Without DBMS any media transmission organization can't think.
- The Database the executive's framework is fundamental for these organizations to store the call subtleties and month to month postpaid bills in the information base.

8. ACCOUNT

- The information base administration framework is utilized for putting away data about deals, holding and acquisition of monetary instruments.

9. ONLINE SHOPPING

- These days, web-based shopping has become a major pattern. Nobody needs to visit the shop and burn through their time.
- Everybody needs to shop through web based shopping sites, (for example, Amazon, Flipkart, Snapdeal) from home.

10. HUMAN RESOURCE MANAGEMENT

- Big firms or organizations have numerous specialists or representatives working under them. They store data about worker's compensation, assessment, and work with the assistance of an information base administration framework (DBMS).

11. MANUFACTURING

- Manufacturing organizations make various kinds of items and deal them consistently.
- To keep the data about their items like bills, acquisition of the item, amount, inventory network the executives, information base administration framework is utilized.

12. AIRLINE RESERVATION SYSTEM

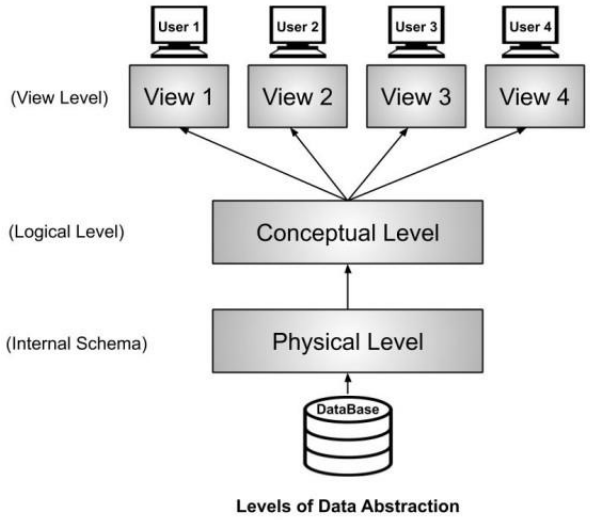
- This framework is equivalent to the railroad reservation framework. This framework additionally utilizes an information base administration framework to store the records of flight takeoff, appearance, and defer status.

VIEWS OF DATA

- The provision given by the database management system to view the data by the people or users of enterprise is called as views of data.
- Indirectly the way how the people of the enterprise look for the data in the database is called as views of data.
- The main function of the database management system is to provide users a concrete view of data. This concrete view is called as abstract view. The meaning of the word abstract here is, 'to hide the details'.

DATA ABSTRACTION

- The database management system developers hide the complexity of data in the system. It is done through several levels of abstraction to simplify user interaction with the system.
- The details of the data are hidden at the THREE levels namely Physical Level, Logical Level and View Level.
- The View Level is at the highest level of data abstraction whereas the Physical Level is the lowest level. In between comes the Logical Level.
- Starting from the lowest level the levels of abstraction details are as follows:



Levels of Data Abstraction

PHYSICAL LEVEL

- This is the lowest level of data abstraction.
- It describes the actual storage of data.
- The complex low-level data structures are detailed in this level.
- This level of data view is normally available only to the Database Administrator.
- The details of this level are hidden normally to the Application Programmers and other inexperienced users of DBMS.
- So, overall, the entire database is described in this level that is physical or internal level. It is a very complex level to understand.

LOGICAL LEVEL

- Next higher level of data abstraction is Logical Level.
- It describes the data stored in the database.
- This level comprises the information that is actually stored in the database in the form of tables.
- It also stores the relationship among the data entities in relatively simple structures

VIEW LEVEL

- This is the highest level of data abstraction. It provides many views for the same database.
- The users at this level need to access only a part of database. So, their interaction with the system is simplified at this level.
- Users view data in the form of rows and columns. Tables and relations are used to store data.
- Multiple views of the same database may exist. Users can just view the data and interact with the database, storage and implementation details are hidden from them.

THREE LEVEL ARCHITECTURE OF DATABASE

- The ANSI-SPARC database architecture is the basis of most of the modern databases.
- The three levels present in this architecture are Physical level, Conceptual level and External level.
- The details of these levels are as follows –

PHYSICAL LEVEL

- This is the lowest level in the three level architecture.
- It is also known as the internal level.
- The physical level describes how data is actually stored in the database.
- In the lowest level, this data is stored in the external hard drives in the form of bits and at a little high level, it can be said that the data is stored in files and folders.
- The physical level also discusses compression and encryption techniques.

CONCEPTUAL LEVEL

- The conceptual level is at a higher level than the physical level. It is also known as the logical level.
- It describes how the database appears to the users conceptually and the relationships between various data tables.
- The conceptual level does not care for how the data in the database is actually stored.

EXTERNAL LEVEL

- This is the highest level in the three level architecture and closest to the user. It is also known as the view level.
- The external level only shows the relevant database content to the users in the form of views and hides the rest of the data.
- So different users can see the database as a different view as per their individual requirements.

DATA INDEPENDENCE

- Data independence means a change of data at one level should not affect another level. Two types of data independence are present in this architecture:

- **PHYSICAL DATA INDEPENDENCE**
  - Any change in the physical location of tables and indexes should not affect the conceptual level or external view of data. This data independence is easy to achieve and implemented by most of the DBMS.

- **CONCEPTUAL DATA INDEPENDENCE**
  - The data at conceptual level schema and external level schema must be independent.
  - This means a change in conceptual schema should not affect external schema.
  - e.g.; Adding or deleting attributes of a table should not affect the user's view of the table.
  - But this type of independence is difficult to achieve as compared to physical data independence because the changes in conceptual schema are reflected in the user's view.

COMPONENTS OF DBMS

- Hardware, Software, Data, Database Access Language, Procedures and Users all together form the components of a DBMS.

- **HARDWARE**
  - The hardware is the actual computer system used for keeping and accessing the database.
  - The conventional DBMS hardware consists of secondary storage devices such as hard disks.
  - Databases run on the range of machines from microcomputers to mainframes.

- **SOFTWARE**
  - Software is the actual DBMS between the physical database and the users of the system.
  - All the requests from the user for accessing the database are handled by DBMS.

- **DATA**
  - It is an important component of the database management system. The main task of DBMS is to process the data.
  - Databases are used to store the data, retrieved, and updated to and from the databases.

- **USERS**
  - **DATA ADMINISTRATOR**
    - Database Administrator (DBA) is a person/team who defines the schema and also controls the 3 levels of database.
    - DBA is also responsible for providing security to the database and he allows only the authorized users to access/modify the database.
  - **DATABASE DESIGNERS**
    - Database Designers are the users who design the structure of database which includes tables, indexes, views, constraints, triggers, stored procedures.
  - **NAÏVE / PARAMETRIC END USERS**
    - Parametric End Users are the unsophisticated who don't have any DBMS knowledge but they frequently use the database applications in their daily life to get the desired results.
    - For examples, Railway's ticket booking users are naive users. Clerks in any bank is a naive user because they don't have any DBMS knowledge but they still use the database and perform their given task.
  - **SOPHISTICATED USERS**
    - Sophisticated users can be engineers, scientists, business analyst, who are familiar with the database. They can develop their own data base applications according to their requirement.
    - They don't write the program code but they interact the database by writing SQL queries directly through the query processor.
  - **APPLICATION PROGRAM**
    - Application Program are the back end programmers who writes the code for the application programs.
    - They are the computer professionals. These programs could be written in Programming languages such as Visual Basic, Developer, C, FORTRAN, COBOL etc.
  - **SYSTEM ANALYST**
    - System Analyst is a user who analyzes the requirements of parametric end users. They check whether all the requirements of end users are satisfied.

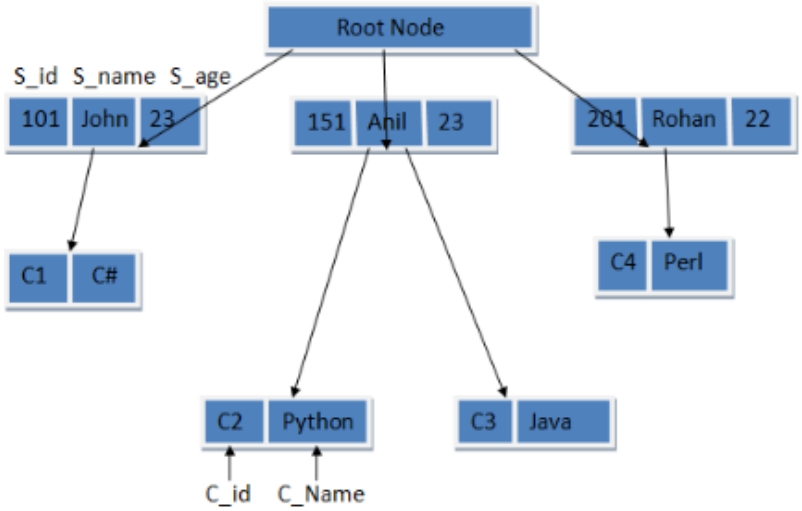
DATA MODEL

- Data Model gives us an idea that how the final system will look like after its complete implementation.
- Data Models are used to show how data is stored, connected, accessed and updated in the database management system.
- Data Models in DBMS are:

- Hierarchical Model
- Network Model
- Entity-Relationship Model
- Relational Model

HIERARCHICAL MODEL

- This model organizes the data in the hierarchical tree structure.
- The hierarchy starts from the root which has root data and then it expands in the form of a tree adding child node to the parent node.



- Features of a Hierarchical Model

1. One-to-many relationship

The data here is organized in a tree-like structure where the one-to-many relationship is between the datatypes. Also, there can be only one path from parent to any node.

2. Parent-Child Relationship

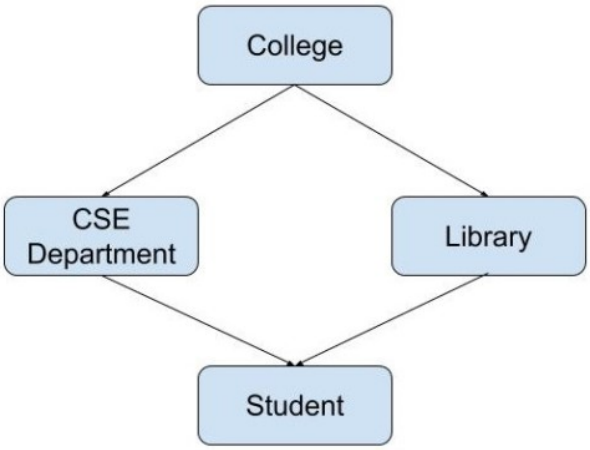
Each child node has a parent node but a parent node can have more than one child node. Multiple parents are not allowed.

3. Deletion Problem

If a parent node is deleted, then the child node is automatically deleted.

NETWORK MODEL

- This model is an extension of the hierarchical model.
- This model is the same as the hierarchical model, the only difference is that a record can have more than one parent.



Network Model

- Features of a Network Model

1. **Ability to Merge more Relationships:**

   This model has the ability to manage one-to-one relationships as well as many-to-many relationships.

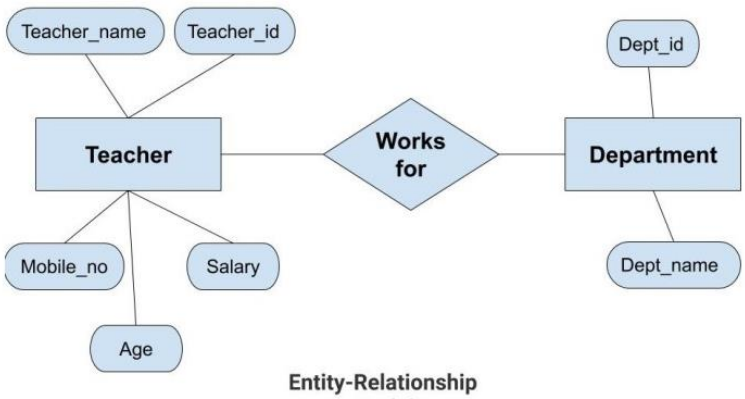
2. **Many paths**

   As there are more relationships so there can be more than one path to the same record. This makes data access fast and simple.

ENTITY-RELATIONSHIP MODEL

- Entity-Relationship Model or simply ER Model is a high-level data model diagram.
- It is also very easy for the developers to understand the system by just looking at the ER diagram.
- ER diagram has the following three components:

1. Entities: Entity is a real-world thing.
2. Attributes: An entity contains a real-world property called attribute.
3. Relationship: Relationship tells how two attributes are related.



Entity-Relationship Model

- Features of ER Model
  1. Graphical Representation for Better Understanding

     It is very easy and simple to understand so it can be used by the developers to communicate with the stakeholders.

  2. Database Design

     This model helps the database designers to build the database and is widely used in database design.

• RELATIONAL MODEL

- Relational Model is the most widely used model.
- In this model, the data is maintained in the form of a two-dimensional table.
- All the information is stored in the form of row and columns.
- The basic structure of a relational model is tables. So, the tables are also called relations in the relational model.

Emp_id	Emp_name	Job_name	Salary	Mobile_no	Dep_id	Project_id
AfterA001	John	Engineer	100000	9111037890	2	99
AfterA002	Adam	Analyst	50000	9587569214	3	100
AfterA003	Kande	Manager	890000	7895212355	2	65

EMPLOYEE TABLE