Geographic Information Systems

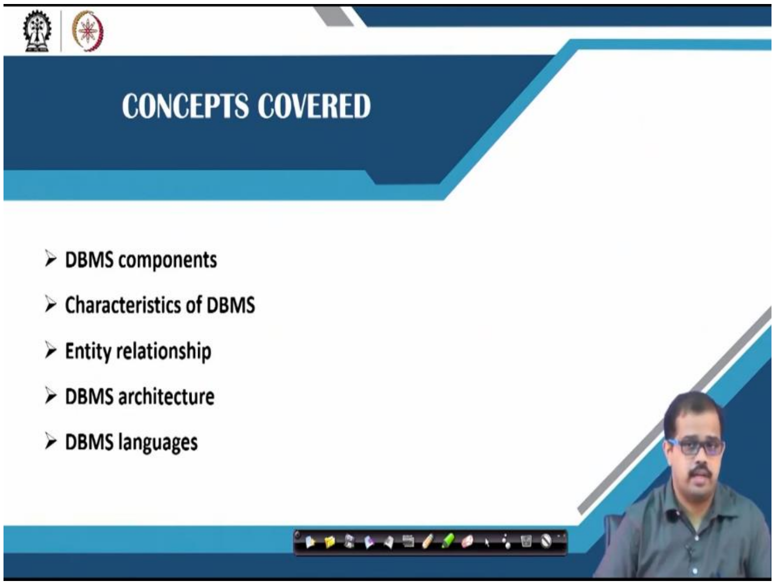
Prof. Bharth H Aithal
Ranbir and Chitra Gupta School of Infrastructure Design and Management
Indian Institute of Technology - Kharagpur

Module No # 07
Lecture No # 32
Database Management System – Introduction

Hello Namaste welcome back to the module where we are looking at how a database works. What is the database management system etce., So in the previous class we learnt what do we mean by a data then we evolved in terms of what do you mean by database then look at the database system and the database management system which is an application. So we have looked at all of these.

Now let us get into what do you mean by database management system. How does it works? What are the different relations that exist in this particular database management system and how we perceive that particular relationship will look at all of these in this particular lecture.

(Refer Slide Time 01:04)



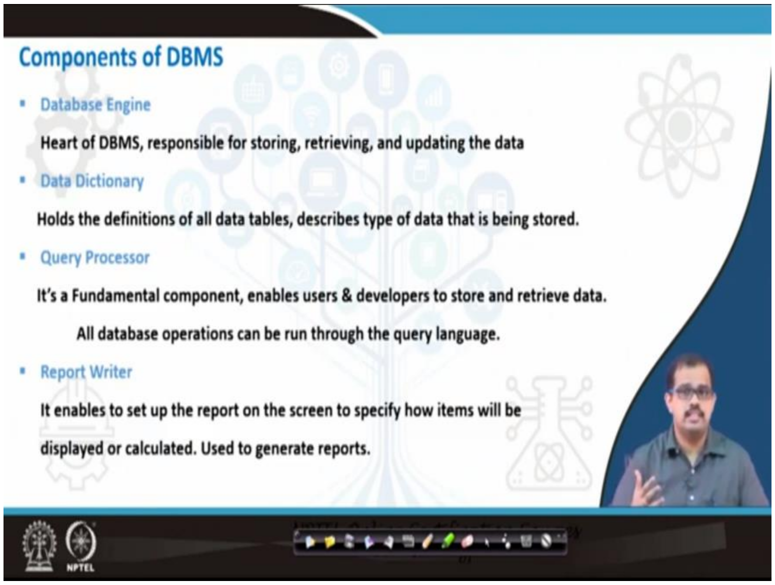
CONCEPTS COVERED

- DBMS components
- Characteristics of DBMS
- Entity relationship
- DBMS architecture
- DBMS languages

And when we look at the concepts that would be covered in this particular class first thing that I would speak about is what do you mean by a database management system component. Components in a database management system. Then characteristics of a database management system then we define what is an entity. So now we will also take it ahead and look at something called as an entity relationship. How entities are related what are the different relationships that exist and the same concepts that we were looking at the spatial aspects. So we would now put it in terms of a database.

Now look at with then we would look at what do you mean by database management system architecture ok. What is the architecture of a database management system? Then there are a database management system language a broad definition of languages I am not getting into very specific a very broad system of different languages in database management system.

(Refer Slide Time 02:03)



Components of DBMS

- Database Engine

  Heart of DBMS, responsible for storing, retrieving, and updating the data
- Data Dictionary

  Holds the definitions of all data tables, describes type of data that is being stored.
- Query Processor

  It's a Fundamental component, enables users & developers to store and retrieve data.

  All database operations can be run through the query language.
- Report Writer

  It enables to set up the report on the screen to specify how items will be displayed or calculated. Used to generate reports.

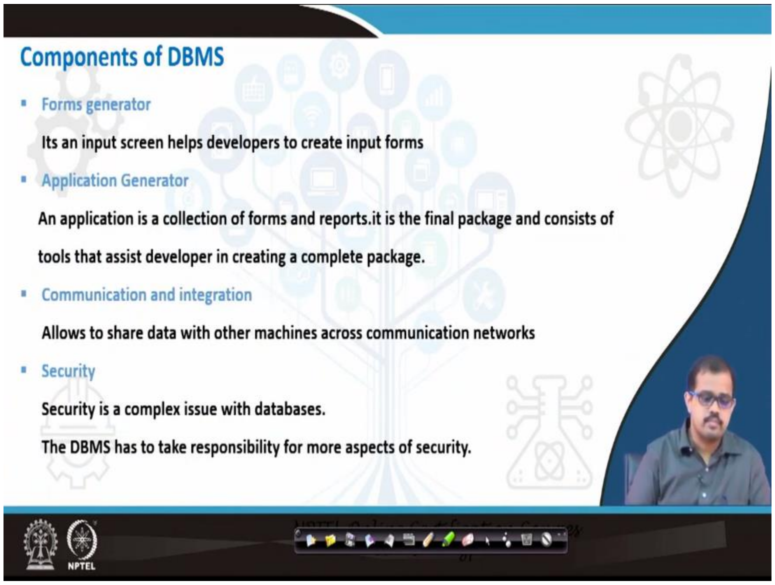
So when we look at the component of the database a first and foremost important component of database is the database engine. When we look at this database engine it is the heart of the database management system responsible for storing, retrieving and updating the data. So this is like a backbone or the heart of a database management system. So this is a this runs through the entire structure and is imp very important for storing, retrieving and updating the data.

There is another concept called data dictionary. So dictionary it hold definition for all the data tables okay. It is like a defining o key in which actually is defining all the tables are at the definition of all the table that describes the type of the data and being that are being stored in that particular database ok. So then we have a query processor so when we look at these towards query and processor which means it is a processing a query. Let us a fun it is basically a fundamental component which enables users and developers to store and retrieve data ok. It is both storing and retrieving.

So when we say querying you are either you are actually storing somewhere the data retrieving the data. So all these database operations can run through query language like a sequential query language SQL, MySQL very well-known query languages the MySQL ok. Though it has a word or a period of time but now also MySQL as quite well known. And when you when the next component is report writer.

So though this particular component is more or less not the evolved over a period of time but it enables us to setup a report to on the screen to specify how items will be displayed or calculated. So these are normally used to generate reports. So in many system many database systems they have different way of defining it but report writer is generally to actually compile all the information and generate report about the those information’s.

(Refer Slide Time 04:22)



Components of DBMS

- Forms generator

  Its an input screen helps developers to create input forms
- Application Generator

  An application is a collection of forms and reports.it is the final package and consists of tools that assist developer in creating a complete package.
- Communication and integration

  Allows to share data with other machines across communication networks
- Security

  Security is a complex issue with databases.

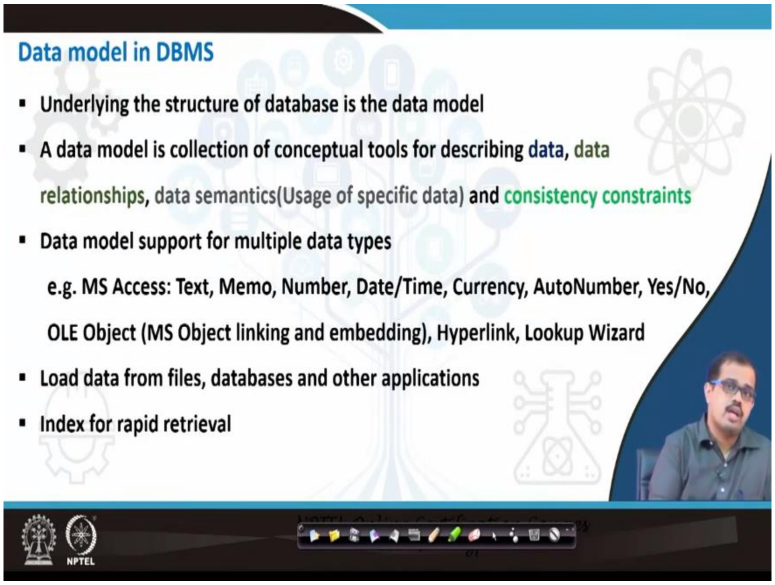
  The DBMS has to take responsibility for more aspects of security.

Then we have something called as a forms generator. It is an input screen that helps developer to create forms. So are it is input screen that helps developers to create and way of connecting into the system. Then you have an application generator and when you look at an application generator as an application is this collection in terms of forms that is this number of forms what would have been generated it is nothing but the collection of those form and reports by the report generator.

It is a final package that consist of tools and that assist developer in creating the complete package. Either a developer or a user to create a complete package. Normally it has done by a developer ok. Then you have communication and integration which allows to share data with other machines across different communication networks. So it is based on relationship it based on addresses it is based on the way you want to query data or you want to access data that in this communication and integration works.

Then the security being a very complex issue in the database. So it has many layers of security many kinds of security that will be embedded in terms of a component in a database management system.

(Refer Slide Time 05:40)



Data model in DBMS

- Underlying the structure of database is the data model
- A data model is collection of conceptual tools for describing data, data relationships, data semantics(Usage of specific data) and consistency constraints
- Data model support for multiple data types
  e.g. MS Access: Text, Memo, Number, Date/Time, Currency, AutoNumber, Yes/No, OLE Object (MS Object linking and embedding), Hyperlink, Lookup Wizard
- Load data from files, databases and other applications
- Index for rapid retrieval

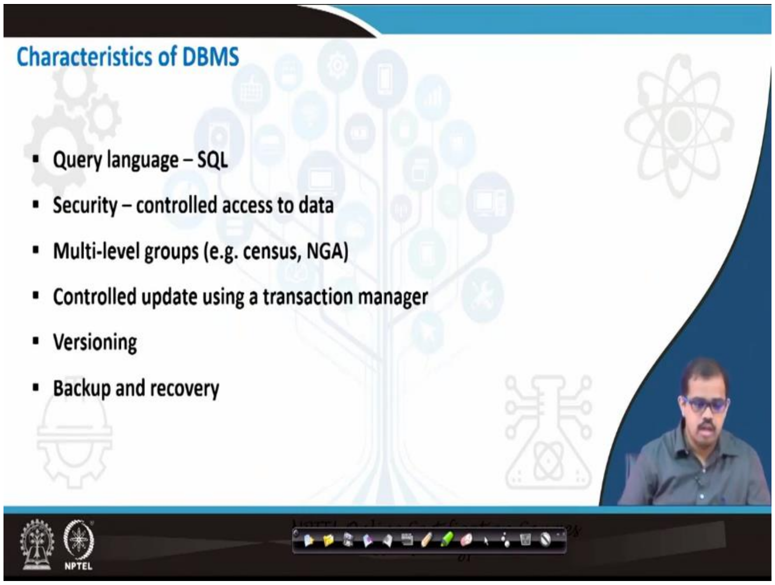
And when you look at database management system always any of those any systems whether it is storing data or integrating data, collecting data and then storing processing etc., always is dependent on the model ok. A kind of model which actually defines how the structure is there how it is saved? How it is retrieved and how different systems work. So when we look at the underlying the structure of the database always you have as I said is a data model. A data model is the collection so let us be very specific here it is a collection of conceptual tools. So it is a collection of conceptual tools that describe data ok.

So when you are speaking about data that are tools that can describe data that can describe data relationship describe data semantic. When I say semantics, these are usage of very specific in data ok. And how the consistent of a consistent consistency a constraint are maintained. So these are very important in terms of a data model. So it is the tool which actually connects or describes all of these ok. That is why that is what is called the data model ok.

So when I say a data model it has support for multiple data types it can have for example you can have MS access and this MS access, text in terms of memo on numbers, date whatever. If you have if you used MS access most of this can be done. Whether it is qualitative, quantitative, ordinary data whatever kind of data that you want to put in can be put in such databases. And when you look at another important aspect is that it is the data model actually starts from the where the data is stored as a file in the database.

Either loading of the databases from different databases and processing it an application this all thing is done using a data model. So if your nutshell if you have to say this is an index for our rapid retrieval ok. So it is like a book you have number of pages in a book. Each page have let us say a certain number of pages have certain chapters in a book. So it acts as a index chapter and helps in retrieving managing and processing any anything in that book is nothing but a data model ok. So it is in a nutshell we can just say it is an index for processing retrieving any information in that particular database is nothing but a data model.

(Refer Slide Time 08:22)



Characteristics of DBMS

- Query language – SQL
- Security – controlled access to data
- Multi-level groups (e.g. census, NGA)
- Controlled update using a transaction manager
- Versioning
- Backup and recovery

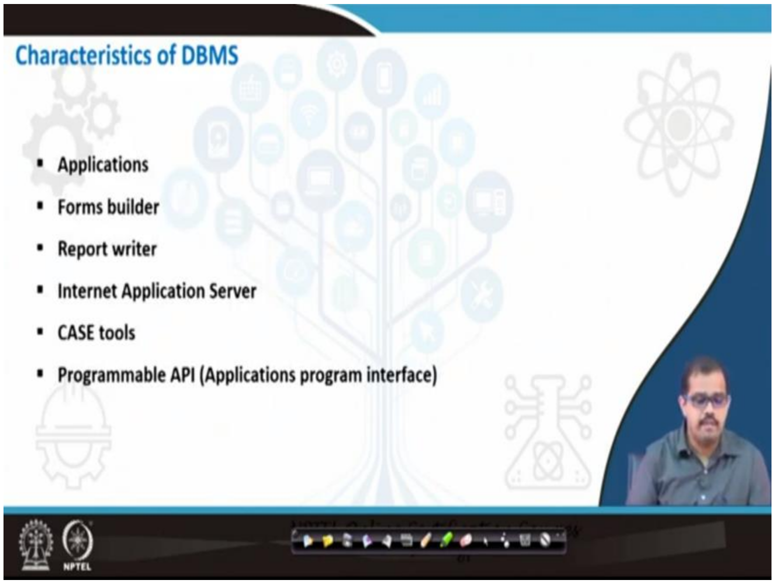
When you look at characteristic of a database management system it consist of query language. For example a sequential query language a MySQL is another query language this is a very intrinsic part of how the query is directed in to the system and that is how that system would program. Then you have security control of the data or the controlled access to the data. So that that is the security forms another characteristic of database management system the better the security the better the database management system multilevel groups for example sensors NGA.

So all of these groups can be easily accessed. So you have different groups for different groups for different access you have different ways of accessing each of this you have different user groups which have different powers of accessing this data. So that is how a database management system is done. Then you have controlled update using a transaction manager. So many of times you may also see that you have to visit a particular site or a visit a particular or write to a particular person in order to have some restriction removed facts of a particular data. So that is what is called controlled update using a transaction manager.

Then you have versioning is a very important form of the database which actually keeps a track of what has improved and how the previous errors or loopholes in the database have been corrected over a period of time. Then the last one is the backup and recovery and this forms a very important part of any database because without backups then maybe certain issues sometimes maybe that particular system where the entire database is cornered on the system that is acting as a server may fail because everything as its own self life.

So once it is self-life is over it may fail. So in order to have a backup if the systematical backup is done now if that database can automatically do it either to a cloud system or to a system that is physically there then or it is remotely connected then it becomes a best part of the database. So it is very intrinsic characteristic of a database. So always most of the database system management system have their own way of backing up the data and also the recovery of data. That is also very essentially very important in term of having database management system.

(Refer slide Time 10:57)

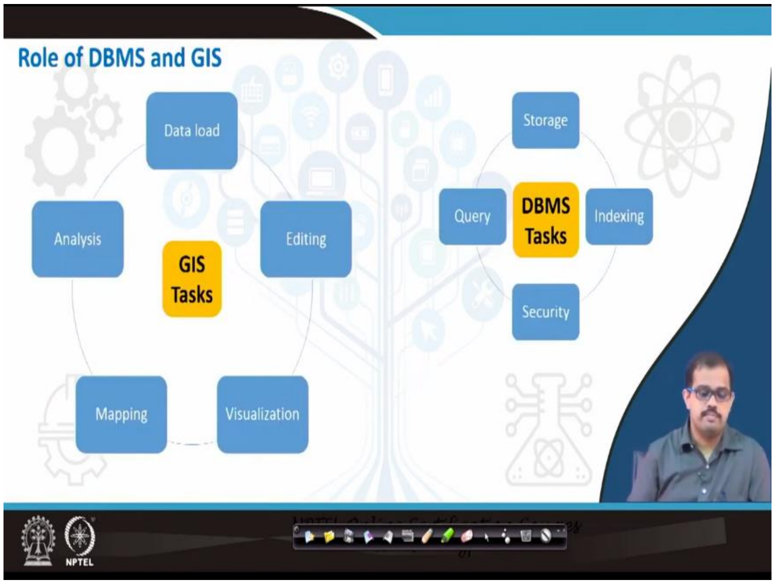


Characteristics of DBMS

- Applications
- Forms builder
- Report writer
- Internet Application Server
- CASE tools
- Programmable API (Applications program interface)

And when we look at the characteristics of database management system we also can look at in terms of application I have already spoke about each of these. This application there are form builders a report writer and internet application server which is actually interacting with the user and different portals. For example if you are looking at censors, if you are looking at Aadhar, if you are looking at any other different portals so it is actually looking at it acts as a bridge between the user and that particular portal.

Then you have case tools which is looking at particular case in terms of the query handling or query processing etc., Then the most important point of today's system is programmable API or applications program interface. So this is has become extremely important in terms of how do we define a API? How do we mange an API? How do we present an API? So that because that is also a very challenging task today that becomes a characteristic of database management system. **(Refer Slide Time 12:05)**



Role of DBMS and GIS

GIS Tasks

Data load

Editing

Visualization

Mapping

Analysis

DBMS Tasks

Storage

Indexing

Security

Query

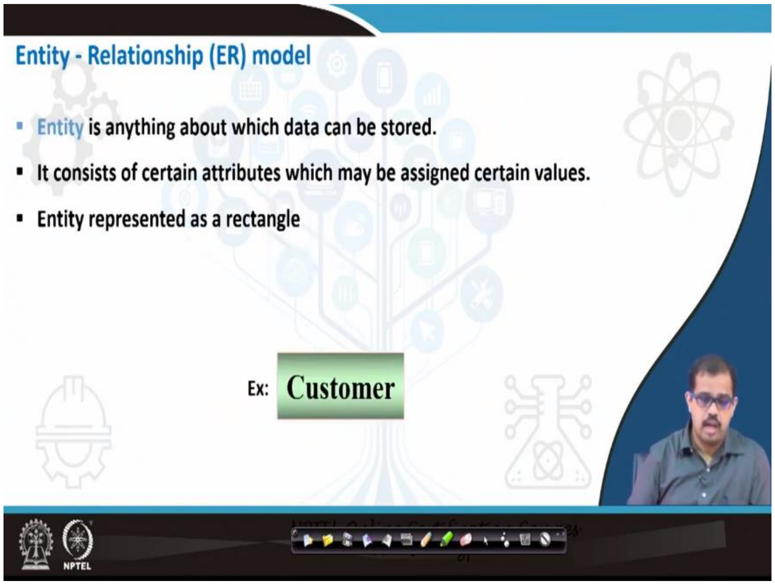
So if you have to put in what is a role of a database management system at GIS? If you look at it when you look at GIS task ok let us take it separately and let us look at it. If you look at GIS task what is it task? The first thing is loading of a data. So when you acquire the data then you load the data. Once you have load the data you edit the data you either edit it or analyze it. So once you have analyze it you can visualize it ok and finally you will map it that is how the GIS task work. So the same task is can be done in terms of a database management system.

So this task is somewhere for example when you are looking at storage the database management task is first storage indexing it then querying it and maintaining it security layer. So when you are integrating both of this is data load. So when you have a data load you are actually looking at storage and security of both of these then looking at editing and then analysis. So this is where the query part comes.

And when you look at the next part again this is where you have visualization again it is part of a query and it is a part of a storage and mapping. Once you have a mapping both indexing and querying comes into effect and when you loot at both of these finally you also have a security layers for different types of understanding or maybe for download maybe for visualization. So what kind of a query it may be dependent on not everything every layer would be available for download but it can be viewed online.

So similarly when integrating both of these database management system and GIS it would have become a very extremely powerful system in terms of land management or any of those management which has spatial features that is embedded into it.

(Refer Slide Time 14:06)



Entity - Relationship (ER) model

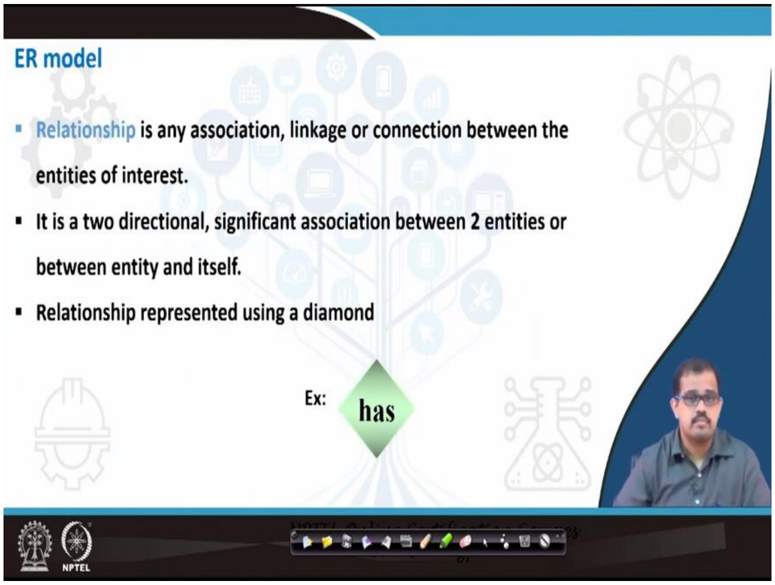
- Entity is anything about which data can be stored.
- It consists of certain attributes which may be assigned certain values.
- Entity represented as a rectangle

And most of our today's database should and must have special features in order to survive in today's context. So we when we go back to some few weeks before we even spoke about an entity right. So when we speak about entity it is a separate portion of the earth surface. Similarly here in the database entity can be anything entity can be just a particular customer ok. For example if you look at the banking database entity is just a customer ok.

So when we define entity and entity is anything that about which a data can be stored ok. So if you if there is lot of data that is being generated if there is a particular task that becomes an entity. If there is a let us say you are you open an account ok. So your name and your name and your address you are an entity basically and based on you transaction what is your age? What is your Aadhar number? What is your pan number? So all of these becomes your different attributes that are stored are the data that is stored in this particular entity ok.

So entity is nothing but nothing is anything that about which a data is to be stored. Then it consists of certain attributes which may be assigned certain values as I said or your id number, you pan number all these are attributes to that entity. Then entity is represented in a form of rectangle in terms of a database. For example I am trying to represent a entity of a customer. So entity customer is nice said in the form of a rectangle.

(Refer Slide Time 15:48)



ER model

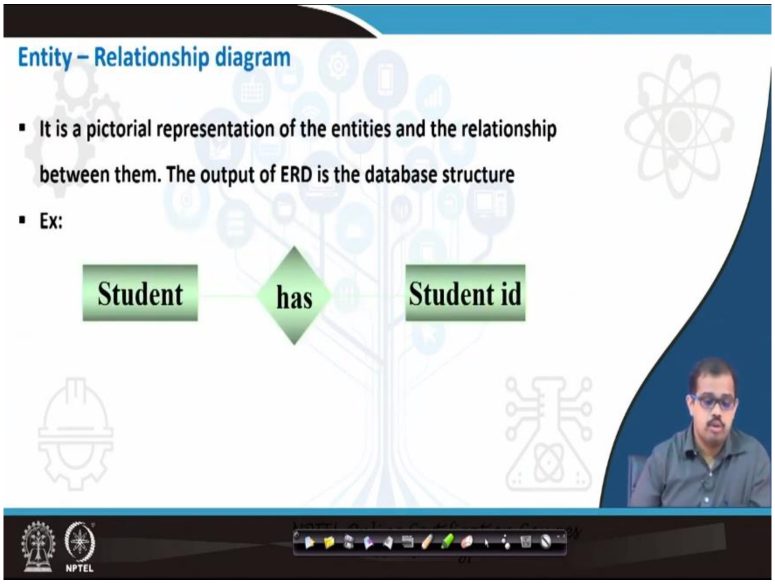
- Relationship is any association, linkage or connection between the entities of interest.
- It is a two directional, significant association between 2 entities or between entity and itself.
- Relationship represented using a diamond

Ex:

Now when we are trying to look at the relationship the ER model. So when I say relation in any association linkages or connection between the entities of interest. Now let us say there is a person who has an account in bank 1 and has an account in bank 2. Now maybe that new ID number or Aadhar number has a linkage for both of these banks ok pan number also can link both of these banks.

And that link can be set in terms of relations person A has a account in this particular bank and also has account in the other bank. So this becomes the relation between the 2 entities 2 banks where the customer is the entity and that particular ER entity has 2 accounts in 2 different banks. It is two directional normally significant association between 2 entities are between entities and itself ok and itself. So you should be very careful when we are representing it here. Normally relationships is represented in form of a diamond. So in the entity is in the form of a rectangle whereas relationship is in the form of a diamond ok.

(Refer Slide Time 17:05)

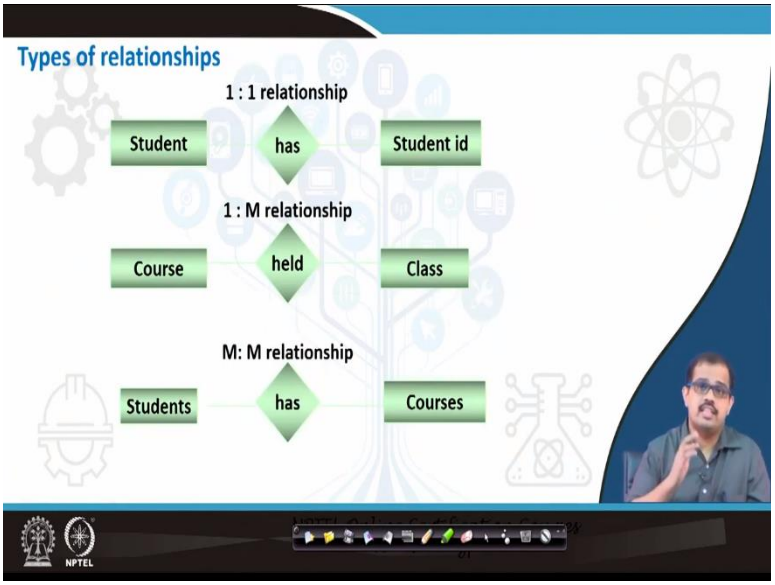


Entity – Relationship diagram

- It is a pictorial representation of the entities and the relationship between them. The output of ERD is the database structure
- Ex:

So when we look at the entity relationship diagram this is how normally an example of an entity relationship. For example a student is a entity here has its relationship and student ID is an again an entity ok. So students maybe any students name he has a particular student id is what we mean ok. So relationship is defined in a form of a diamond. Whereas the entities are in the form of a rectangle. So this this is how a entity relationship is built and entity relationship diagram is developed.

(Refer Slide Time 17:46)

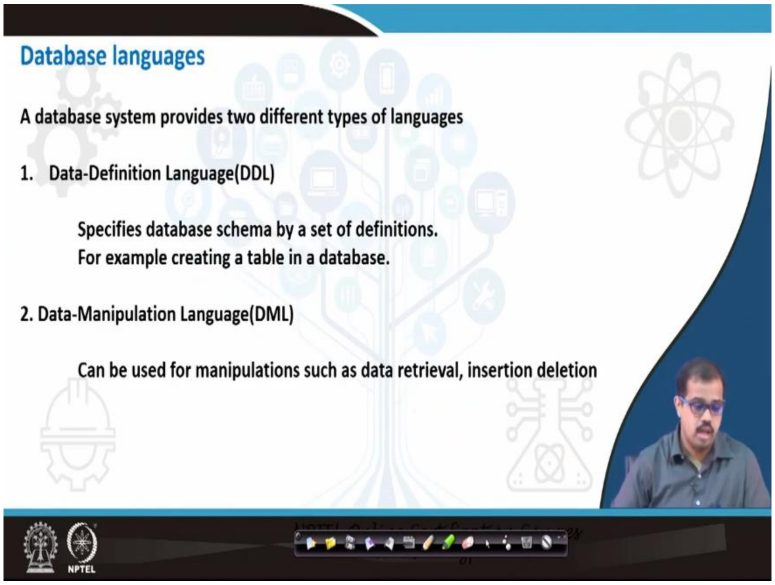


Types of relationships

And when we look at the types of entity we have 1 is to 1 relationship which means there is one particular entity and one particular relation that is being defined you have one is to many. For example there is a particular class or course ok. Let us say GIS is a course is being held in this particular classes ok which means to say that you may have many classes ok many sessions many modules but the this particular course is a single course ok. So it is one entity with many relationship many with many entities. Now you have many to many also you have multiple students taking a multiple courses.

So that means that it is many students has many different courses or have considered many different courses. So this becomes 3 different types of relationship. So the first kind of relationship has 1 is to 1 relationship which is actually saying that there is one entity one relation is related to the other entity ok. Whereas one is to many is defined between one entity and many entities a relationship that connects. Then you have many to many relationship where you have many entities and many entities that are related through various means. So that is about the types of relationship that exists.

(Refer Slide Time 19:12)



Database languages

A database system provides two different types of languages

1. Data-Definition Language(DDL)

   Specifies database schema by a set of definitions.
   For example creating a table in a database.

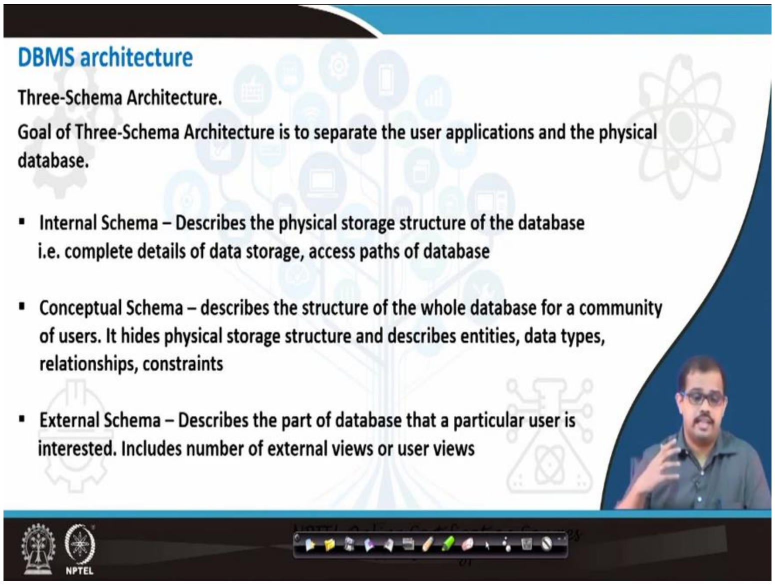
2. Data-Manipulation Language(DML)

   Can be used for manipulations such as data retrieval, insertion deletion

So as I said database language there are various ways of representing a database languages but when you broadly define the database languages the database systems are provide two times two different types of languages. So the one is the data definition language it specifies database schema by a set of definitions ok. For example creating a table in the database is the way of defining a particular database. How the particular table is being created?

So that is a database definition language. We have data manipulation language which can be used for manipulation such as data retrieval, insertion, deletion all of these are human quantification and providing qualitative information all of these are data manipulations. So this is handled by the data manipulation language ok. So when I say language it has series of syntax and semantics associated with it that is why it is called language. So we have a data definition language, we have a data manipulation language ok. So both of these languages together becomes account of a database language ok.

(Refer Slide Time 20:27)



DBMS architecture

Three-Schema Architecture.

Goal of Three-Schema Architecture is to separate the user applications and the physical database.

- Internal Schema – Describes the physical storage structure of the database i.e. complete details of data storage, access paths of database
- Conceptual Schema – describes the structure of the whole database for a community of users. It hides physical storage structure and describes entities, data types, relationships, constraints
- External Schema – Describes the part of database that a particular user is interested. Includes number of external views or user views

So the next thing we would look at is a database management system architecture. So when we look at a database management system architecture we look at this for a 3 tire system ok. When you look at the architecture the very important aspect of defining the 3 schema architecture is to separate the user application with the physical databases. So you have user application itself is different whereas your physical application a physical interface itself is completely different or the database is completely different.

So you have to separate both of these that is why a three time schema is always in used in terms of a DBMS architecture. Though it has evolved over the period of time today you can fine more complex schemas that are available in terms of defining the architecture of a database management system. But when we speak about basic schema this is how the database management system architecture is defined.

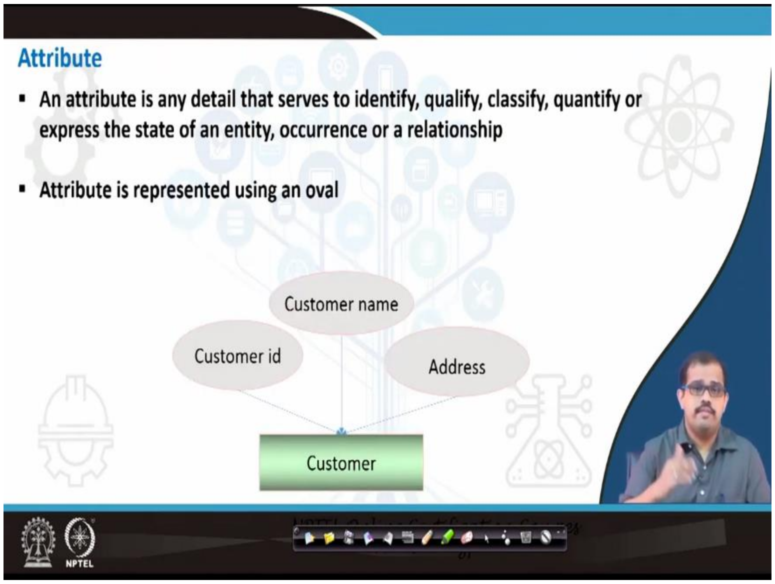
The first one is the internal schema. When we look at the internal schema this describes a physical storage structure of the database that is complete details of a data storage access path of a database. So this actually says that where the physical storage has been done ok which part of the system. Where physically the entire data stored ok and what is the structure of that particular database ok it also tell you how to access these different data that have been stored in different paths in terms of path storage or to the database.

So how do we access how do it stored and where it is stored? So all of these are provided in the internal schema ok. When we look at the next part of the 3-tier system it is called as a conceptual schema. Conceptual schema describes the structure of whole database for a community of users. It hides physical storage structures and describes entities datatypes and relationships and constraints.

For example if you go into a website it does not give you what are different ways of store your data has been stored. Where it has been stored? What are the kinds of data that is available? But it gives you a lot of drop down menus maybe or picking boxes. So where in you just select some of these and use it for your analysis. So that is nothing but your concept conceptual schema. Where you do not know how the physical storage has been done but you only know how do you how what are the different data's are available?

How do you access it what are the querying mechanism and how do you use such a data? Then the third one you have is that external schema. External schema describes a part of a database that that a particular user is interested in or particular user needs to understand and includes number of external views or user views. And when you look at the external schema this is schema that can be defined based on the user base and can be interacted in terms of how the user needs it to be manipulated or maybe even modified.

(Refer Slide Time 23:42)



Attribute

- An attribute is any detail that serves to identify, qualify, classify, quantify or express the state of an entity, occurrence or a relationship
- Attribute is represented using an oval

So once you have understood what is the architecture the next thing that comes here is an attribute. So whenever we look at any database so you have attribute as a very important factor that is defining the entire database. Otherwise you do not have database that is existent if you had just have a customer name in the entire database it is now rather than called a ledger than a database.

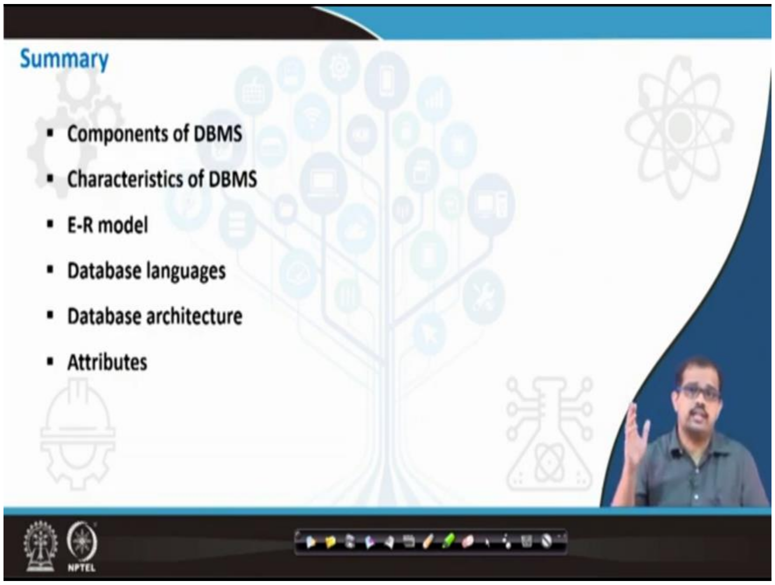
Database should always have an attribute information. When I say attribute information these can be any detail it depends on what kind of system? What kind of user? What kind of instance you are trying to use a particular database. So it can be any detail that serves. So to identify, qualify, quantify or classify or express the state of entity. So entity is a very important aspect so and then it can be on and off it may be colors it may be your different types.

So all of these are defining the identity qualifier quality classify, quantify so all of these are an entity that can be expressed occurring in terms of a relationship that is nothing but an attribute. Attribute and entity are always connected through a relationship. And this relationship is defined based on what kind of information that entity is providing ok or different entities are providing. Then when look at this attribute is represented using an oval basically.

For example when we look at it here where as I previously said the rectangle is representation of an entity. Now entity is a customer now customer has an address, customer has a customer ID, customer has a customer name, customer can have Aadhar number, pan number, customer can also have different details that is required for a like an maybe an address, alternate address, the mobile phone number which is a very important part.

So and your and where the branch is located so all of these information are attribute information’s. So this is what is another part of a database management system. So we have 3 kinds of boxes one is rectangle, other one is diamond and another one is oval. So oval represents attributes, rectangle is represented by an for an entity. And whereas relationships are designed based on your diamond boxes.

(Refer Slide Time 26:15)



Summary

- Components of DBMS
- Characteristics of DBMS
- E-R model
- Database languages
- Database architecture
- Attributes

So in order to summarize this particular class we first look at the components of a database. We clearly define what are different quantum components of the database though this has been previously discussed. But we looked at what are the different components of a database and how it can be of used to us. Then we looked at what a characteristic of database. Characteristic is extremely important for any database management system to and to be understood. If you have not understood what are the characteristic of a database management system?

What type of database management system and what is its characteristic and probably implementation of a complex database managing capability may not be there. So looking at that the characteristic is extremely important. And when we look at the next we looked at the entity relationship model we did define the entity as a soul model which is actually defining any of those information in the data or which is which has data which is connected with data. That is nothing but a entity.

Entity is normally represented in the form of a rectangular boxes whereas relationship is the one that actually defines or connects an entity to different attributes or different entities or different ways of maintaining the database. So that is nothing but a relationship so it has normally in the form of a diamond box and this is nothing but the ER model. And then we look at the data base languages we said that broadly it can be classified into two one is the data definition language other one is a data manipulation language.

So at the we also looked at the database architecture what is the architecture of a database. We also looked at GIS and a database how different parts how a GIS and the database management system works? We looked at how if we take a GIS how the loading of the data? How the data is then converted into different features? How the query is done and finally you have process information.

Similarly when you look at database management system this this particular system works including the security. And when we combine both of these this can become a powerful engine in terms of any applications. So then we looked at finally at attributes. Attributes are again the part of a database that we maintain every data has every database has a major attribute. So when we look at attributes it can be anything that is linked to the entity and most of these attributes are located in our representative of a oval.

So this is what we have learnt in this particular class. So in the next class we will continue this understand and we will go further probably we will I will also introduce to terms called as normalization etc., So which is of extreme importance. So we will look all those more concepts about what do you mean by database management system? What is a database different types of the database in the next class also till then have a nice time thank you.